BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 9, 2003, 11:39:33 ; Search time 37 Seconds
(without alignments)
1813.343 Million cell updates/sec

Title: US-09-687-609A-1
Perfect score: 1371
Sequence: 1 GSHMIEGYECQPIFLNVLEA.....SVQVPKILSGKVKPIYFHTQ 260

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1349.5	98.4	896	6	Q9GKL7	Q9gkl7 sus scrofa
2	1346.5	98.2	912	6	Q8MIK0	Q8mik0 crocuta cro
3	1336.5	97.5	906	4	Q9UN21	Q9un21 homo sapien
4	1330	97.0	730	4	Q13771	Q13771 homo sapien
5	1326.5	96.8	895	6	Q9GKN9	Q9gkn9 sus scrofa
6	1285	93.7	333	13	Q8JH77	Q8jh77 poephila gu
7	1224	89.3	790	13	P70048	P70048 xenopus lae
8	1217	88.8	344	13	Q91445	Q91445 serinus can
9	1061.5	77.4	303	6	O97684	O97684 ovis aries
10	935	68.2	853	13	O93245	O93245 oncorhynchu
11	930	67.8	839	13	Q8QFV2	Q8qfv2 carassius a
12	928	67.7	769	13	O93497	O93497 pagrus majo
13	925	67.5	797	13	Q9PWG5	Q9pwg5 anguilla ja
14	911	66.4	563	13	Q9DDJ4	Q9ddj4 halichoeres
15	909	66.3	848	13	Q9YGV9	Q9ygv9 anguilla ja
16	909	66.3	854	13	O93244	O93244 oncorhynchu
17	896	65.4	763	13	Q8QFV7	Q8qfv7 haplochromi
18	884	64.5	763	13	Q8UWB7	Q8uwb7 oreochromis
19	797	58.1	688	13	Q8UWB8	Q8uwb8 oreochromis
20	792	57.8	232	13	Q9IA30	Q9ia30 anolis caro
21	782	57.0	692	13	Q9W6F4	Q9w6f4 haplochromi
22	774	56.5	703	13	Q8UVY3	Q8uvy3 xenopus lae
23	773	56.4	732	13	Q9DDU9	Q9ddu9 xenopus lae
24	768	56.0	583	13	Q9DEV4	Q9dev4 xenopus lae
25	766	55.9	939	6	Q9GLW0	Q9glw0 canis famil
26	759	55.4	698	11	Q8BW69	Q8bw69 mus musculu
27	759	55.4	711	13	Q8AYI2	Q8ayi2 rana dybows
28	734	53.5	710	13	Q9IBD5	Q9ibd5 anguilla ja
29	720	52.5	360	13	O42274	O42274 crocodylus
30	720	52.5	438	13	Q90ZM7	Q90zm7 petromyzon
31	711	51.9	348	13	Q91425	Q91425 cnemidophor
32	678	49.5	401	13	Q90ZM6	Q90zm6 petromyzon
33	677.5	49.4	359	13	Q9IAC6	Q9iac6 oncorhynchu
34	673	49.1	298	6	Q28547	Q28547 ovis aries
35	668	48.7	993	13	Q8JJ89	Q8jj89 haplochromi
36	664.5	48.5	779	13	Q8JJ92	Q8jj92 haplochromi
37	662	48.3	689	13	Q8QGX5	Q8qgx5 anguilla ja
38	662	48.3	982	6	Q9N0W8	Q9n0w8 saimiri sci
39	658.5	48.0	793	13	Q8JJ91	Q8jj91 haplochromi
40	658.5	48.0	802	13	Q8JJ90	Q8jj90 haplochromi
41	654.5	47.7	793	11	Q8R5J0	Q8r5j0 rattus norv
42	654.5	47.7	794	11	Q8R463	Q8r463 rattus norv
43	628	45.8	284	13	Q90Y00	Q90y00 petromyzon
44	596	43.5	831	4	Q8TDS3	Q8tds3 homo sapien
45	590	43.0	258	6	Q9BDJ7	Q9bdj7 ovis aries

ALIGNMENTS

RESULT 1
Q9GKL7
ID Q9GKL7 PRELIMINARY; PRT; 896 AA.
AC Q9GKL7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Androgen receptor.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary;
RA Song J.H., Fahrenkrug S.C., Rohrer G.A., Wise T.H., Ford J.J.;
RT "Sus scrofa androgen receptor (AR) coding sequence.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary;
RA Song J.H., Fahrenkrug S.C., Rohrer G.A., Wise T.H., Ford J.J.;
RT "Porcine androgen receptor (AR) cDNA cloning: Expression in pituitary
RT associates with FSH secretion in boars.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
DR EMBL; AF202775; AAG37994.1; -.
DR HSSP; P06536; 1GDC.
DR InterPro; IPR001103; Andrgn_receptor.
DR InterPro; IPR000536; Hormone_rec_lig.
DR InterPro; IPR001628; Znf_C4steroid.
DR Pfam; PF02166; Androgen_recep; 1.
DR Pfam; PF00104; hormone_rec; 1.
DR Pfam; PF00105; zf-C4; 1.
DR PRINTS; PR00047; STROIDFINGER.
DR ProDom; PD000035; Znf_C4steroid; 1.
DR SMART; SM00430; HOLI; 1.
DR SMART; SM00399; ZnF_C4; 1.

```
DR   PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW   DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
KW   Transcription regulation; Zinc; Zinc-finger.
SQ   SEQUENCE   896 AA;  97094 MW;  440F5F6E73BDC796 CRC64;

  Query Match           98.4%;  Score 1349.5;  DB 6;  Length 896;
  Best Local Similarity   99.6%;  Pred. No. 3.5e-119;
  Matches  258;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 697

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        698 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 757

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 817

Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        818 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 877

Qy        242 VQVPKILSGKVKPIYFHTQ 260
              ||||||||||||||||||
Db        878 VQVPKILSGKVKPIYFHTQ 896


RESULT 2
Q8MIK0
ID   Q8MIK0      PRELIMINARY;      PRT;   912 AA.
AC   Q8MIK0;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Androgen receptor.
OS   Crocuta crocuta (Spotted hyena).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Hyaenidae; Crocuta.
OX   NCBI_TaxID=9678;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Catalano S., Avila D.M., McPhaul M.J., Marsico S., Wilson J.D.,
RA   Glickman S.E.;
RT   "Absence of association of the virilization of the female spotted
RT   hyena with alterations of the amino acid sequence of the androgen
RT   receptor (AR).";
RL   Mol. Cell. Endocrinol. 0:0-0(2002).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
DR   EMBL; AY128705; AAM96904.1; -.
DR   InterPro; IPR001103; Andrgn_receptor.
DR   InterPro; IPR000536; Hormone_rec_lig.
DR   InterPro; IPR001628; Znf_C4steroid.
DR   Pfam; PF02166; Androgen_recep; 1.
DR   Pfam; PF00104; hormone_rec; 1.
DR   Pfam; PF00105; zf-C4; 1.
DR   PRINTS; PR00047; STROIDFINGER.
DR   ProDom; PD000035; Znf_C4steroid; 1.
DR   SMART; SM00430; HOLI; 1.
DR   SMART; SM00399; ZnF_C4; 1.
DR   PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW   DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
KW   Transcription regulation; Zinc; Zinc-finger.
SQ   SEQUENCE   912 AA;  99957 MW;  8F5EC35B1E743C51 CRC64;

  Query Match           98.2%;  Score 1346.5;  DB 6;  Length 912;
  Best Local Similarity   99.2%;  Pred. No. 6.8e-119;
  Matches  257;  Conservative    1;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 713

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        714 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 773

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db        774 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKDLDR 833

Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        834 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 893

Qy        242 VQVPKILSGKVKPIYFHTQ 260
              ||||||||||||||||||
Db        894 VQVPKILSGKVKPIYFHTQ 912


RESULT 3
Q9UN21
ID   Q9UN21      PRELIMINARY;      PRT;   906 AA.
AC   Q9UN21;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Androgen receptor.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Breast carcinoma;
RA   Jin C.H., Urcan-Bisel M.S., Schrader W.T.;
RT   "Androgen receptor sequences in human mammary carcinoma MDA-MB-453
RT   cells.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
DR   EMBL; AF162704; AAD45921.1; -.
DR   HSSP; P06536; 1GDC.
DR   InterPro; IPR001103; Andrgn_receptor.
DR   InterPro; IPR000536; Hormone_rec_lig.
DR   InterPro; IPR001628; Znf_C4steroid.
DR   Pfam; PF02166; Androgen_recep; 1.
DR   Pfam; PF00104; hormone_rec; 1.
DR   Pfam; PF00105; zf-C4; 1.
DR   PRINTS; PR00047; STROIDFINGER.
DR   ProDom; PD000035; Znf_C4steroid; 1.
DR   SMART; SM00430; HOLI; 1.
DR   SMART; SM00399; ZnF_C4; 1.
DR   PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW   DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
KW   Transcription regulation; Zinc; Zinc-finger.
SQ   SEQUENCE   906 AA;  97884 MW;  0A1FA8802B2EDDAF CRC64;

  Query Match           97.5%;  Score 1336.5;  DB 4;  Length 906;
  Best Local Similarity   98.5%;  Pred. No. 6e-118;
  Matches  255;  Conservative    1;  Mismatches    2;  Indels    1;  Gaps    1;

Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        649 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 707

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        708 ALPGLRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 767

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              ||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db        768 YSQCVRMRHLSQEFGWLQITPQEFLCMKAMLLFSIIPVDGLKNQKFFDELRMNYIKELDR 827
```

```
Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              ||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db        828 IIACKRKNPTSCSRRFYQLTKLLDSVHPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 887

Qy        242 VQVPKILSGKVKPIYFHTQ 260
              ||||||||||||||||||
Db        888 VQVPKILSGKVKPIYFHTQ 906


RESULT 4
Q13771
ID   Q13771      PRELIMINARY;      PRT;   730 AA.
AC   Q13771;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Androgen receptor.
GN   AR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RX   MEDLINE=90258935; PubMed=2342476;
RA   Govindan M.V.;
RT   "Specific region in hormone binding domain is essential for hormone
RT   binding and trans-activation by human androgen receptor.";
RL   Mol. Endocrinol. 4:417-427(1990).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
DR   EMBL; M73069; AAA51735.1; -.
DR   HSSP; P06536; 1GDC.
DR   InterPro; IPR001103; Andrgn_receptor.
DR   InterPro; IPR000536; Hormone_rec_lig.
DR   InterPro; IPR001628; Znf_C4steroid.
DR   Pfam; PF02166; Androgen_recep; 1.
DR   Pfam; PF00104; hormone_rec; 1.
DR   Pfam; PF00105; zf-C4; 1.
DR   PRINTS; PR00047; STROIDPINGER.
DR   ProDom; PD000035; Znf_C4steroid; 1.
DR   SMART; SM00430; HOLI; 1.
DR   SMART; SM00399; ZnF_C4; 1.
DR   PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW   DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
KW   Transcription regulation; Zinc; Zinc-finger.
SQ   SEQUENCE   730 AA;  79150 MW;  918B47C3B41B80C9 CRC64;

  Query Match             97.0%;  Score 1330;  DB 4;  Length 730;
  Best Local Similarity   96.6%;  Pred. No. 1.9e-117;
  Matches  257;  Conservative    0;  Mismatches    1;  Indels    8;  Gaps    2;

Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        466 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 524

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        525 ALPGLRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 584

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKA-------LLLFSIIPVDGLKNQKFFDELRMN 174
              |||||||||||||||||||||||||||       |||||||||||||||||||||||||
Db        585 YSQCVRMRHLSQEFGWLQITPQEFLCMKAMLIFFFFLLLFSIIPVDGLKNQKFFDELRMN 644

Qy        175 YIKELDRIIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPE 234
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        645 YIKELDRIIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPE 704

Qy        235 MMAEIISVQVPKILSGKVKPIYFHTQ 260
              |||||||||||||||||||||||||
Db        705 MMAEIISVQVPKILSGKVKPIYFHTQ 730
```

```
RESULT 5
Q9GKN9
ID   Q9GKN9      PRELIMINARY;      PRT;   895 AA.
AC   Q9GKN9;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Androgen receptor AR.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20539123; PubMed=11086548;
RA   Trakooljul N., Ponsuksili S., Schellander K., Wimmers K.;
RT   "A highly polymorphic repetitive polypyrimidine/polypurine (CCTTT)n
RT   sequence in the 5' untranslated sequence of the porcine androgen
RT   receptor gene.";
RL   Anim. Genet. 31:288-289(2000).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
DR   EMBL; AF161717; AAG40566.1; -.
DR   HSSP; P06536; 1GDC.
DR   InterPro; IPR001103; Andrgn_receptor.
DR   InterPro; IPR000536; Hormone_rec_lig.
DR   InterPro; IPR001628; Znf_C4steroid.
DR   Pfam; PF02166; Androgen_recep; 1.
DR   Pfam; PF00104; hormone_rec; 1.
DR   Pfam; PF00105; zf-C4; 1.
DR   PRINTS; PR00047; STROIDPINGER.
DR   ProDom; PD000035; Znf_C4steroid; 1.
DR   SMART; SM00430; HOLI; 1.
DR   SMART; SM00399; ZnF_C4; 1.
DR   PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW   DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
KW   Transcription regulation; Zinc; Zinc-finger.
SQ   SEQUENCE   895 AA;  97156 MW;  923C2FDD1F7E4779 CRC64;

  Query Match             96.8%;  Score 1326.5;  DB 6;  Length 895;
  Best Local Similarity   98.1%;  Pred. No. 5.2e-117;
  Matches  254;  Conservative    1;  Mismatches    3;  Indels    1;  Gaps    1;

Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        638 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 696

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        697 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 756

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        757 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 816

Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              || : |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        817 IIVMQEKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 876

Qy        242 VQVPKILSGKVKPIYFHTQ 260
              ||||||||||||||||||
Db        877 VQVPKILSGKVKPIYFHTQ 895


RESULT 6
Q8JH77
ID   Q8JH77      PRELIMINARY;      PRT;   333 AA.
AC   Q8JH77;
DT   01-OCT-2002 (TrEMBLrel. 22; Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
```

```
DE   Androgen receptor (Fragment).
OS   Poephila guttata (Zebra finch) (Taeniopygia guttata).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Passeriformes; Estrildidae;
OC   Estrildinae; Taeniopygia.
OX   NCBI_TaxID=59729;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Perlman W.R., Ramachandran B., Arnold A.P.;
RT   "Expression of Androgen Receptor mRNA in the Late Embryonic and Early
RT   Post-hatch Zebra Finch Brain.";
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
DR   EMBL; AF532914; AAM96699.1; -.
DR   InterPro; IPR000536; Hormone_rec_lig.
DR   InterPro; IPR001723; Stdhrmn_receptor.
DR   InterPro; IPR001628; Znf_C4steroid.
DR   Pfam; PF00104; hormone_rec; 1.
DR   Pfam; PF00105; zf-C4; 1.
DR   PRINTS; PR00398; STRDHORMONER.
DR   ProDom; PD000035; Znf_C4steroid; 1.
DR   SMART; SM00430; HOLI; 1.
DR   SMART; SM00399; ZnF_C4; 1.
KW   DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
KW   Transcription regulation; Zinc; Zinc-finger.
FT   NON_TER       1       1
SQ   SEQUENCE   333 AA;  38536 MW;  B36A7778B710B790 CRC64;

  Query Match             93.7%;  Score 1285;  DB 13;  Length 333;
  Best Local Similarity   93.4%;  Pred. No. 1.4e-113;
  Matches  239;  Conservative   12;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          5 IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALP 64
              |:|||||||||||||||||||||||||||:|||||: |||||||||||||||:|||||||||
Db         78 IDGYECQPIFLNVLEAIEPGVVCAGHDNSQPDSFSNLLSSLNELGERQLVYVVKWAKALP 137

Qy         65 GFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQ 124
              ||||||||||::|||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 GFRNLHVDDQMSIIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQ 197

Qy        125 CVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIA 184
              |:|||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db        198 CIRMRHLSQEFGWLQITPQEFLCMKALLFFSIIPVDGLKNQKLFDELRMNYIKELDRIIA 257

Qy        185 CKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQV 244
              ||||||||||||||||:|||| |||::|||||||||||:||||||:||||||||||||||
Db        258 CKRKNPTSCSRRFYQLTKVLDSVHPIAKDLHQFTFDLLIKAHMVSVDYPEMMAEIISVQV 317

Qy        245 PKILSGKVKPIYFHTQ 260
              |||||||||||||| |
Db        318 PKILSGKVKPIYFHAQ 333


RESULT 7
P70048
ID   P70048      PRELIMINARY;      PRT;   790 AA.
AC   P70048;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-JAN-1999 (TrEMBLrel. 09, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Androgen receptor alpha isoform.
GN   XL ALPHA AR.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93376782; PubMed=7690145;
RA   Fischer L., Catz D., Kelley D.;
RT   "An androgen receptor mRNA isoform associated with hormone-induced
RT   cell proliferation.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:8254-8258(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95324753; PubMed=7601302;
RA   Fischer L.M., Catz D., Kelley D.B.;
RT   "Androgen-directed development of the Xenopus laevis larynx: control
RT   of androgen receptor expression and tissue differentiation.";
RL   Dev. Biol. 170:115-126(1995).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Kelley D.B., Kamenetz F.R., Kelley D.B., Badea T.C.;
RL   Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
DR   EMBL; U67129; AAC97386.1; -.
DR   HSSP; P06536; 1GDC.
DR   InterPro; IPR001103; Andrgn_receptor.
DR   InterPro; IPR000536; Hormone_rec_lig.
DR   InterPro; IPR001628; Znf_C4steroid.
DR   Pfam; PF02166; Androgen_recep; 1.
DR   Pfam; PF00104; hormone_rec; 1.
DR   Pfam; PF00105; zf-C4; 1.
DR   PRINTS; PR00047; STROIDFINGER.
DR   ProDom; PD000035; Znf_C4steroid; 1.
DR   SMART; SM00430; HOLI; 1.
DR   SMART; SM00399; ZnF_C4; 1.
DR   PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW   DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
KW   Transcription regulation; Zinc; Zinc-finger.
SQ   SEQUENCE   790 AA;  86973 MW;  9E7FC136CCCC3906 CRC64;

  Query Match             89.3%;  Score 1224;  DB 13;  Length 790;
  Best Local Similarity   88.4%;  Pred. No. 2.4e-107;
  Matches  228;  Conservative   15;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          1 GSHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWA 60
              |   :||| |||||||||||||| |||||||||||||| ||||||||||||||||||||||
Db        531 GIPQLEGYSCQPIFLNVLEAIEPVVVCAGHDNNQPDSFALLLSSLNELGERQLVHVVKWA 590

Qy         61 KALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSR 120
              |||||||||| ||| ||||||||:|||||||| ||||||||||||||||||||||||||||
Db        591 KALPGFRNLHVSDQMTVIQYSWMGLMIFAMGWRSFKNVNSRMLYFAPDLVFNEYRMHKSR 650

Qy        121 MYSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELD 180
              |||||||:|||||||||||||:|||||||||||||||:|||:|| ||||||||||||
Db        651 MYSQCVRLRHLSQEFGWLQITPEEFLCMKALLLFSIIPVEGLKDQKCFDELRMNYIKELD 710

Qy        181 RIIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEII 240
              |:|:||| || | ||||:||||||||||||||||||||| :|: |||||:||||:|||
Db        711 RVISCKRNNPASSSRRFFQLTKLLDSVQPIARELHQFTFDLFVKAQMVSVDYPEMMSEII 770

Qy        241 SVQVPKILSGKVKPIYFH 258
              ||||||||||:|||:|||
Db        771 SVQVPKILSGRVKPLYFH 788


RESULT 8
Q91445
ID   Q91445      PRELIMINARY;      PRT;   344 AA.
AC   Q91445;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Androgen receptor (Fragment).
GN   AR.
OS   Serinus canaria (Canary).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Passeriformes; Passeroidea;
OC   Fringillidae; Carduelinae; Serinus.
OX   NCBI_TaxID=9135;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RC   TISSUE=Testis;
RX   MEDLINE=94130808; PubMed=8299561;
RA   Nastiuk K.L., Clayton D.F.;
RT   "Seasonal and tissue-specific regulation of canary androgen receptor
RT   messenger ribonucleic acid.";
RL   Endocrinology 134:640-649(1994).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
DR   EMBL; L25901; AAA17402.1; -.
DR   HSSP; P06536; 1GLU.
DR   InterPro; IPR000536; Hormone_rec_lig.
DR   InterPro; IPR001723; Stdhrmn_receptor.
DR   InterPro; IPR001628; Znf_C4steroid.
DR   Pfam; PF00104; hormone_rec; 1.
DR   Pfam; PF00105; zf-C4; 1.
DR   PRINTS; PR00398; STRDHORMONER.
DR   PRINTS; PR00047; STROIDFINGER.
DR   ProDom; PD000035; Znf_C4steroid; 1.
DR   SMART; SM00430; HOLI; 1.
DR   SMART; SM00399; ZnF_C4; 1.
KW   DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
KW   Transcription regulation; Zinc; Zinc-finger.
FT   NON_TER       1      1
FT   NON_TER     344    344
SQ   SEQUENCE   344 AA;  39376 MW;  713676394FC0B030 CRC64;

  Query Match             88.8%;  Score 1217;  DB 13;  Length 344;
  Best Local Similarity   92.7%;  Pred. No. 4e-107;
  Matches  227;  Conservative   13;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          5 IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALP 64
              |:|||||||||||||||||||||||||:|||||: ||:|||||||||||:|||||||||
Db        100 IDGYECQPIFLNVLEAIEPGVVCAGHDNSQPDSFSNLLTSLNELGERQLVYVVKWAKALP 159

Qy         65 GFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQ 124
              ||||||||||::|||||||||||||||||||||||||||||||||||||||||||||||
Db        160 GFRNLHVDDQMSIIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQ 219

Qy        125 CVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIA 184
              |:|||||||||||||||| |||||||| |||||||||||||| ||||||||||||||||||
Db        220 CIRMRHLSQEFGWLQITPQGFLCMKALLFFSIIPVDGLKNQKLFDELRMNYIKELDRIIA 279

Qy        185 CKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQV 244
              |||||||||||||||||:|||| |||::||||||||||:||||||:||||||||||||||
Db        280 CKRKNPTSCSRRFYQLTKVLDSVTPIAKDLHQFTFDLLIKAHMVSVDYPEMMAEIISVQV 339

Qy        245 PKILS 249
              |||||
Db        340 PKILS 344


RESULT 9
O97684
ID   O97684      PRELIMINARY;      PRT;   303 AA.
AC   O97684;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Androgen receptor (Fragment).
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Dorset;
RA   Thonney M.L., Ignotz G., Richards H.M.;
RT   "Partial sequence of an ovine androgen receptor cDNA.";
RL   Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
DR   EMBL; AF105713; AAC97958.1; -.
DR   HSSP; P06536; 1GLU.
DR   InterPro; IPR000536; Hormone_rec_lig.
DR   InterPro; IPR001723; Stdhrmn_receptor.
DR   InterPro; IPR001628; Znf_C4steroid.
DR   Pfam; PF00104; hormone_rec; 1.
DR   Pfam; PF00105; zf-C4; 1.
DR   PRINTS; PR00398; STRDHORMONER.
DR   PRINTS; PR00047; STROIDFINGER.
DR   ProDom; PD000035; Znf_C4steroid; 1.
DR   SMART; SM00430; HOLI; 1.
DR   SMART; SM00399; ZnF_C4; 1.
KW   DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
KW   Transcription regulation; Zinc; Zinc-finger.
FT   NON_TER       1      1
FT   NON_TER     303    303
SQ   SEQUENCE   303 AA;  34673 MW;  D14D3950C7DC69D1 CRC64;

  Query Match             77.4%;  Score 1061.5;  DB 6;  Length 303;
  Best Local Similarity   99.5%;  Pred. No. 1.9e-92;
  Matches  201;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        103 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 161

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        162 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 221

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        222 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 281

Qy        182 IIACKRKNPTSCSRRFYQLTKL 203
              ||||||||||||||||||||||
Db        282 IIACKRKNPTSCSRRFYQLTKL 303


RESULT 10
O93245
ID   O93245      PRELIMINARY;      PRT;   853 AA.
AC   O93245;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Androgen receptor beta.
GN   AR-BETA.
OS   Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX   NCBI_TaxID=8022;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99150354; PubMed=10026186;
RA   Takeo J., Yamashita S.;
RT   "Two distinct isoforms of cDNA encoding rainbow trout androgen
RT   receptors.";
RL   J. Biol. Chem. 274:5674-5680(1999).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
DR   EMBL; AB012096; BAA32785.1; -.
DR   HSSP; P06536; 1GDC.
DR   InterPro; IPR001103; Andrgn_receptor.
DR   InterPro; IPR000536; Hormone_rec_lig.
DR   InterPro; IPR001723; Stdhrmn_receptor.
DR   InterPro; IPR001628; Znf_C4steroid.
DR   Pfam; PF02166; Androgen_recep; 1.
DR   Pfam; PF00104; hormone_rec; 1.
DR   Pfam; PF00105; zf-C4; 1.
DR   PRINTS; PR00398; STRDHORMONER.
DR   PRINTS; PR00047; STROIDFINGER.
DR   ProDom; PD000035; Znf_C4steroid; 1.
```

```
DR   SMART; SM00430; HOLI; 1.
DR   SMART; SM00399; ZnF_C4; 1.
DR   PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW   DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
KW   Transcription regulation; Zinc; Zinc-finger.
SQ   SEQUENCE   853 AA;  95776 MW;  65EFF5FD3B36F4C4 CRC64;

  Query Match           68.2%;  Score 935;  DB 13;  Length 853;
  Best Local Similarity  70.4%;  Pred. No. 6.6e-80;
  Matches  178;  Conservative   35;  Mismatches   36;  Indels    4;  Gaps    2;

Qy          8 YECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFR 67
              :  | :|||:||:||| || |||| |||| | ||:|||||||||||| |||||| :||||
Db        602 FHSQLVFLNILESIEPEVVNAGHDHCQPDSAAVLLTSLNELGERQLVKVVKWAKGMPGFR 661

Qy         68 NLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVR 127
              |||||||| |||:||||:||| :|||| |||:||||||||||||::||| | |:  |:|
Db        662 NLHVDDQMTVIQHSWMGVMVFGLGWRSYKNVNARMLYFAPDLVFNDHRMHISSMFEHCIR 721

Qy        128 MRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKR 187
              ||||||:|  ||:| :|||||||||||||||||:||:|||||: || ||||:|    |
Db        722 MRHLSQQFVLLQVTQEEFLCMKALLLFSIIPVDGLKSQKYFDELRLTYINELDRVINYGR 781

Qy        188 KNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMV--SVDFPEMMAEIISVQVP 245
              |  ::||:||||||:|:||:||| |:| |||||| |::  :   | ||||:|||||| ||
Db        782 K--SNCSQRFYQLTRLMDSLQPIVRKLQQFTFDLFIQAQSLPTKVSFPEMIAEIISVHVP 839

Qy        246 KILSGKVKPIYFH 258
              |||:|  ||| ||
Db        840 KILAGLAKPILFH 852


RESULT 11
Q8QFV2
ID   Q8QFV2      PRELIMINARY;      PRT;   839 AA.
AC   Q8QFV2;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Androgen receptor.
GN   AR.
OS   Carassius auratus (Goldfish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Carassius.
OX   NCBI_TaxID=7957;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Betka M., Rothberg S.C., Callard G.V.;
RT   "Carassius auratus Androgen Receptor.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
DR   EMBL; AY090897; AAM09278.1; -.
DR   InterPro; IPR001103; Andrgn_receptor.
DR   InterPro; IPR000536; Hormone_rec_lig.
DR   InterPro; IPR001628; Znf_C4steroid.
DR   Pfam; PF02166; Androgen_recep; 1.
DR   Pfam; PF00104; hormone_rec; 1.
DR   Pfam; PF00105; zf-C4; 1.
DR   ProDom; PD000035; Znf_C4steroid; 1.
DR   SMART; SM00430; HOLI; 1.
DR   SMART; SM00399; ZnF_C4; 1.
DR   PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW   DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
KW   Transcription regulation; Zinc; Zinc-finger.
SQ   SEQUENCE   839 AA;  93168 MW;  A534DC3169C0B1F4 CRC64;

  Query Match           67.8%;  Score 930;  DB 13;  Length 839;
  Best Local Similarity  70.4%;  Pred. No. 1.9e-79;
  Matches  178;  Conservative   35;  Mismatches   36;  Indels    4;  Gaps    2;

Qy          8 YECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFR 67
              :  | ||||:||:||| || |||| |||| |||:|||||||||| |||||| |||||
Db        588 FHSQLIFLNILESIEPEVVNAGHDHAQPDSAVALLTSLNELGERQLVKVVKWAKGLPGFR 647

Qy         68 NLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVR 127
              ||||||| |||::|||:||||:||||: | |:|||||||||||::||| | ||  ||:
Db        648 NLHVDDQMTVIQHTWMGVMVFALGWRSYKNANARMLYFAPDLVFNDHRMHISSMYEHCVQ 707

Qy        128 MRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKR 187
              |:||||||  ||:| :|||||||||||||||:|||:||:||: || ||||:|    |
Db        708 MKHLSQEFVLLQVTQEEFLCMKALLLFSIIPVEGLKSQKYFDDLRLTYINELDRLINYGR 767

Qy        188 KNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMV--SVDFPEMMAEIISVQVP 245
              |  |:|: || |||:|:||:||| |:||||||| :::  :   | ||||:||||||||
Db        768 K--TNCAMRFQQLTRLMDSLQPIVRKLHQFTFDLFVQARSLPTKVSFPEMIAEIISVQVP 825

Qy        246 KILSGKVKPIYFH 258
              |||:|  ||| ||
Db        826 KILAGLSKPILFH 838


RESULT 12
O93497
ID   O93497      PRELIMINARY;      PRT;   769 AA.
AC   O93497;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Androgen receptor.
OS   Pagrus major (Red sea bream) (Chrysophrys major).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidei;
OC   Sparidae; Pagrus.
OX   NCBI_TaxID=143350;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99326178; PubMed=10395960;
RA   Touhata K., Kinoshita M., Tokuda Y., Toyohara H., Sakaguchi M.,
RA   Yokoyama Y., Yamashita S.;
RT   "Sequence and expression of a cDNA encoding the red seabream androgen
RT   receptor.";
RL   Biochim. Biophys. Acta 1450:481-485(1999).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
DR   EMBL; AB017158; BAA33451.1; -.
DR   HSSP; P06536; 1RGD.
DR   InterPro; IPR001103; Andrgn_receptor.
DR   InterPro; IPR000536; Hormone_rec_lig.
DR   InterPro; IPR001723; Stdhrmn_receptor.
DR   InterPro; IPR001628; Znf_C4steroid.
DR   Pfam; PF02166; Androgen_recep; 1.
DR   Pfam; PF00104; hormone_rec; 1.
DR   Pfam; PF00105; zf-C4; 1.
DR   PRINTS; PR00398; STRDHORMONER.
DR   PRINTS; PR00047; STROIDFINGER.
DR   ProDom; PD000035; Znf_C4steroid; 1.
DR   SMART; SM00430; HOLI; 1.
DR   SMART; SM00399; ZnF_C4; 1.
DR   PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW   DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
KW   Transcription regulation; Zinc; Zinc-finger.
SQ   SEQUENCE   769 AA;  86081 MW;  9305C627DBE42062 CRC64;

  Query Match           67.7%;  Score 928;  DB 13;  Length 769;
  Best Local Similarity  70.8%;  Pred. No. 2.7e-79;
  Matches  179;  Conservative   29;  Mismatches   41;  Indels    4;  Gaps    2;

Qy          8 YECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFR 67
              :  | :|||||:||| || |||| |||| | ||:|||||||||| |||||| |||||
Db        518 FNSQVVFLNVLESIEPEVVNAGHDYGQPDSAATLLTSLNELGERQLVKVVKWAKGLPGFR 577

Qy         68 NLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVR 127
```

```
                ||||||||| |||:||||:|||  :||||: ||| ||||||||||||||:||| | ||  |:|
Db        578 NLHVDDQMTVIQHSWMGVMVFGLGWRSYKNVNGRMLYFAPDLVFNEHRMHISTMYEHCIR 637

Qy        128 MRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKR 187
                ||||||||  |||| :|||||||||||||||:|||:||:|||||: || ||||:|  :
Db        638 MRHLSQEFLLLQITQEEFLCMKALLLFSIIPVEGLKSQKYFDELRLTYINELDRLINYRM 697

Qy        188 KNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMV--SVDFPEMMAEIISVQVP 245
                  |:||:|||||||:||||:|   ::|||||||| :::  :   | ||||: ||||| ||
Db        698 N--TNCSQRFYQLTRLLDSLQMTVKKLHQFTFDLFVQAQSLPTKVSFPEMIGEIISVHVP 755

Qy        246 KILSGKVKPIYFH 258
                |||:|  ||| ||
Db        756 KILAGLAKPILFH 768


RESULT 13
Q9PWG5
ID   Q9PWG5      PRELIMINARY;      PRT;   797 AA.
AC   Q9PWG5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Androgen receptor-beta.
GN   AR-BETA.
OS   Anguilla japonica (Japanese eel).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;
OC   Anguilla.
OX   NCBI_TaxID=7937;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RX   MEDLINE=99395076; PubMed=10464240;
RA   Ikeuchi T., Todo T., Kobayashi T., Nagahama Y.;
RT   "cDNA cloning of a novel androgen receptor subtype.";
RL   J. Biol. Chem. 274:25205-25205(1999).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
DR   EMBL; AB025361; BAA83805.1; -.
DR   HSSP; P06536; 1GDC.
DR   InterPro; IPR001103; Andrgn_receptor.
DR   InterPro; IPR000536; Hormone_rec_lig.
DR   InterPro; IPR001723; Stdhrmn_receptor.
DR   InterPro; IPR001628; Znf_C4steroid.
DR   Pfam; PF02166; Androgen_recep; 1.
DR   Pfam; PF00104; hormone_rec; 1.
DR   Pfam; PF00105; zf-C4; 1.
DR   PRINTS; PR00398; STRDHORMONER.
DR   PRINTS; PR00047; STROIDFINGER.
DR   ProDom; PD000035; Znf_C4steroid; 1.
DR   SMART; SM00430; HOLI; 1.
DR   SMART; SM00399; ZnF_C4; 1.
DR   PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW   DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
KW   Transcription regulation; Zinc; Zinc-finger.
SQ   SEQUENCE   797 AA;  89924 MW;  CF1D5F21EEEA5691 CRC64;

  Query Match             67.5%;  Score 925;  DB 13;  Length 797;
  Best Local Similarity   69.2%;  Pred. No. 5.4e-79;
  Matches  175;  Conservative   35;  Mismatches   39;  Indels    4;  Gaps    2;

Qy          8 YECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFR 67
              :  | :|||:||:||| || |||| | || | ||:||||||||||  |||||| |||||
Db        546 FPTQSVFLNILESIEPEVVNAGHDYGQTDSAATLLTSLNELGERQLFKVVKWAKGLPGFR 605

Qy         68 NLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVR 127
              |:|||||  |||::|||:||||:||||: |||:||||||||||: ||   | ||  |:|
Db        606 NMHVDDQMTVIQHAWMGVMVFALGWRSYKNVNARMLYFAPDLVFNDRRMRVSSMYEHCIR 665

Qy        128 MRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKR 187
              |||:||||  ||:| |||||||||||||||:||||||:||:||  || ||||:| | |
Db        666 MRHMSQEFVLLQVTHQEFLCMKALLLFSIIPVEGLKNQKYFDDLRNTYINELDRLINCSR 725

Qy        188 KNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMV--SVDFPEMMAEIISVQVP 245
              |  |:||:|| |||:|:||:||| ::|||||||| :::  :   |:||||:|||||| ||
Db        726 K--TNCSQRFIQLTRLMDSLQPIVKKLHQFTFDLFVQAQSLHTKVNFPEMIAEIISVHVP 783

Qy        246 KILSGKVKPIYFH 258
              :||:|  ||| ||
Db        784 RILAGMAKPILFH 796


RESULT 14
Q9DDJ4
ID   Q9DDJ4      PRELIMINARY;      PRT;   563 AA.
AC   Q9DDJ4;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Androgen receptor (Fragment).
GN   AR.
OS   Halichoeres trimaculatus (Three-spot wrasse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidei;
OC   Labridae; Halichoeres.
OX   NCBI_TaxID=147232;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kim S.J., Kei O., Takemura A., Nakamura M.;
RT   "Partial sequence and expression of androgen and estrogen receptor
RT   genes in the protogynous wrasse, Halichoeres trimaculatus.";
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
DR   EMBL; AF326200; AAG48340.1; -.
DR   HSSP; P06536; 1RGD.
DR   InterPro; IPR000536; Hormone_rec_lig.
DR   InterPro; IPR001723; Stdhrmn_receptor.
DR   InterPro; IPR001628; Znf_C4steroid.
DR   Pfam; PF00104; hormone_rec; 1.
DR   Pfam; PF00105; zf-C4; 1.
DR   PRINTS; PR00398; STRDHORMONER.
DR   PRINTS; PR00047; STROIDFINGER.
DR   ProDom; PD000035; Znf_C4steroid; 1.
DR   SMART; SM00430; HOLI; 1.
DR   SMART; SM00399; ZnF_C4; 1.
DR   PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW   DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
KW   Transcription regulation; Zinc; Zinc-finger.
FT   NON_TER      1      1
SQ   SEQUENCE   563 AA;  64352 MW;  A744C3728F004AF6 CRC64;

  Query Match             66.4%;  Score 911;  DB 13;  Length 563;
  Best Local Similarity   69.6%;  Pred. No. 7.5e-78;
  Matches  176;  Conservative   31;  Mismatches   42;  Indels    4;  Gaps    2;

Qy          8 YECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFR 67
              :  | :|||:||:||| || |||| :|||| | ||:|||||||||||| |||||| |||||
Db        312 FNTQMAFLNILESIEPEVVNAGHDYSQPDSAATLLTSLNELGERQLVKVVKWAKGLPGFR 371

Qy         68 NLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVR 127
              |||||| | ||| ||||:||||:||||: ||||||||||||||||:||| | ||  |:|
Db        372 NLHVDDPMTVIQQSWMGVMVFALGWRSYKNVNSRMLYFAPDLVFNEHRMHISTMYEHCIR 431

Qy        128 MRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKR 187
              |||:||||  |||| :|||||||||||||:||:|||:||:||||||: || ||||:|  :
Db        432 MRHISQEFLMLQITQEEFLCMKALLLFSILPVEGLKSQKYFDELRLAYINELDRLINNRM 491

Qy        188 KNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMV--SVDFPEMMAEIISVQVP 245
              |   :||:||| ||:||||:|   ::|||||||| :::  :   | ||||: ||||| ||
Db        492 K--INCSQRFYHLTRLLDSLQMTVKKLHQFTFDLFVQAQSLPTKVSFPEMIGEIISVHVP 549

Qy        246 KILSGKVKPIYFH 258
              |||:|  ||| ||
```

```
Db        550 KILAGLAKPILFH 562


RESULT 15
Q9YGV9
ID   Q9YGV9      PRELIMINARY;      PRT;   848 AA.
AC   Q9YGV9;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Androgen receptor alpha.
OS   Anguilla japonica (Japanese eel).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;
OC   Anguilla.
OX   NCBI_TaxID=7937;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RX   MEDLINE=99119319; PubMed=9918846;
RA   Todo T., Ikeuchi T., Kobayashi T., Nagahama Y.;
RT   "Fish androgen receptor: cDNA cloning, steroid activation of
RT   transcription in transfected mammalian cells, and tissue mRNA
RT   levels.";
RL   Biochem. Biophys. Res. Commun. 254:378-383(1999).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
DR   EMBL; AB023960; BAA75464.1; -.
DR   HSSP; P06536; 1GDC.
DR   InterPro; IPR001103; Andrgn_receptor.
DR   InterPro; IPR000536; Hormone_rec_lig.
DR   InterPro; IPR001723; Stdhrmn_receptor.
DR   InterPro; IPR001628; Znf_C4steroid.
DR   Pfam; PF02166; Androgen_recep; 1.
DR   Pfam; PF00104; hormone_rec; 1.
DR   Pfam; PF00105; zf-C4; 1.
DR   PRINTS; PR00398; STRDHORMONER.
DR   PRINTS; PR00047; STROIDFINGER.
DR   ProDom; PD000035; Znf_C4steroid; 1.
DR   SMART; SM00430; HOLI; 1.
DR   SMART; SM00399; ZnF_C4; 1.
DR   PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW   DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
KW   Transcription regulation; Zinc; Zinc-finger.
SQ   SEQUENCE   848 AA;  94692 MW;  A8889AF7F2E50D3E CRC64;

  Query Match          66.3%;  Score 909;  DB 13;  Length 848;
  Best Local Similarity   68.5%;  Pred. No. 1.9e-77;
  Matches  174;  Conservative   36;  Mismatches   40;  Indels    4;  Gaps    2;

Qy          7 GYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGF 66
              |:  | :|||:|||||| || ||||  |||| |:||:|||||||||||| |||||| :|||
Db        596 GFHTQSMFLNILEAIEPEVVNAGHDYGQPDSAASLLTSLNELGERQLVKVVKWAKGMPGF 655

Qy         67 RNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCV 126
              |:|:||||| |||:||| :||||:||||| || ||||||||||||||:||  | ||  |:
Db        656 RSLYVDDQMTVIQHSWMAVMVFALGWRSFKNVKSRMLYFAPDLVFNEHRMQVSTMYEHCI 715

Qy        127 RMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACK 186
              ||:: ||||  ||:: :||||||||||| |||:||| | |||||| :|| ||||::: :
Db        716 RMKNFSQEFAMLQVSQEEFLCMKALLLFSTIPVEGLKGQNFFDELRRSYINELDRLVSFR 775

Qy        187 RKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVS--VDFPEMMAEIISVQV 244
              |  :||| || |||:||||:||: ::|||||||| ::|  :|  | ||||::||||| |
Db        776 SK--SSCSERFQQLTRLLDSLQPVLKKLHQFTFDLFVQSQNLSNQVCFPEMISEIISVHV 833

Qy        245 PKILSGKVKPIYFH 258
              ||||:| |||| ||
Db        834 PKILAGTVKPILFH 847


Search completed: October  9, 2003, 11:42:15
Job time : 39 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 9, 2003, 11:39:08 ; Search time 11 Seconds
(without alignments)
1111.540 Million cell updates/sec

Title: US-09-687-609A-1
Perfect score: 1371
Sequence: 1 GSHMIEGYECQPIFLNVLEA.....SVQVPKILSGKVKPIYFHTQ 260

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1349.5	98.4	709	1	ANDR_RABIT	P49699 oryctolagus
2	1349.5	98.4	884	1	ANDR_EULFC	O97776 eulemur ful
3	1349.5	98.4	895	1	ANDR_MACFA	O97952 macaca fasc
4	1349.5	98.4	895	1	ANDR_PAPHA	O97960 papio hamad
5	1349.5	98.4	899	1	ANDR_MOUSE	P19091 mus musculu
6	1349.5	98.4	902	1	ANDR_RAT	P15207 rattus norv
7	1349.5	98.4	907	1	ANDR_CANFA	Q9tt90 canis famil
8	1349.5	98.4	911	1	ANDR_PANTR	O97775 pan troglod
9	1349.5	98.4	919	1	ANDR_HUMAN	P10275 homo sapien
10	766	55.9	930	1	PRGR_RABIT	P06186 oryctolagus
11	765	55.8	933	1	PRGR_HUMAN	P06401 homo sapien
12	759	55.4	923	1	PRGR_RAT	Q63449 rattus norv
13	758	55.3	923	1	PRGR_MOUSE	Q00175 mus musculu
14	752	54.9	786	1	PRGR_CHICK	P07812 gallus gall
15	738	53.8	377	1	PRGR_SHEEP	Q28590 ovis aries
16	677.5	49.4	359	1	MCR_ONCMY	Q9iac6 oncorhynchu
17	666	48.6	978	1	MCR_MOUSE	Q8vii8 mus musculu
18	665	48.5	984	1	MCR_HUMAN	P08235 homo sapien
19	664.5	48.5	777	1	GCR_AOTNA	P79686 aotus nancy
20	664.5	48.5	777	1	GCR_SAIBB	O13186 saimiri bol
21	664.5	48.5	778	1	GCR_SAISC	O46567 saimiri sci
22	662.5	48.3	777	1	GCR_HUMAN	P04150 homo sapien
23	662	48.3	981	1	MCR_RAT	P22199 rattus norv
24	662	48.3	982	1	MCR_SAISC	Q9n0w8 saimiri sci
25	661.5	48.2	777	1	GCR_SAGOE	P79269 saguinus oe
26	660.5	48.2	772	1	GCR_RABIT	P59667 oryctolagus
27	660	48.1	612	1	MCR_XENLA	Q91573 xenopus lae
28	657.5	48.0	783	1	GCR_MOUSE	P06537 mus musculu
29	652.5	47.6	776	1	GCR_TUPGB	Q95267 tupaia glis
30	651.5	47.5	807	1	GCR_PAROL	O73673 paralichthy
31	650.5	47.4	776	1	GCR_XENLA	P49844 xenopus lae
32	648.5	47.3	758	1	GCR_ONCMY	P49843 oncorhynchu
33	644.5	47.0	771	1	GCR_CAVPO	P49115 cavia porce
34	638.5	46.6	795	1	GCR_RAT	P06536 rattus norv
35	615	44.9	977	1	MCR_TUPGB	Q29131 tupaia glis
36	590	43.0	258	1	MCR_SHEEP	Q9bdj7 ovis aries
37	580	42.3	711	1	GCR_PIG	Q9n1u3 sus scrofa
38	391	28.5	164	1	MCR_PIG	P79404 sus scrofa
39	387	28.2	180	1	PRGR_MACEU	P79373 macropus eu
40	342	24.9	147	1	MCR_CHICK	Q8qhi2 gallus gall
41	237	17.3	569	1	ESR1_BRARE	P57717 brachydanio
42	228.5	16.7	581	1	ESR1_PAGMA	O42132 pagrus majo
43	227.5	16.6	617	1	ESR1_ICTPU	Q9yhz7 ictalurus p
44	223	16.3	307	1	ESR1_CNEUN	Q91424 cnemidophor
45	222	16.2	535	1	ESR1_SALSA	P50242 salmo salar

ALIGNMENTS

RESULT 1
ANDR_RABIT
ID ANDR_RABIT STANDARD; PRT; 709 AA.
AC P49699;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Androgen receptor (Dihydrotestosterone receptor) (Fragment).
GN AR OR NR3C4.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand white; TISSUE=Prostate;
RX MEDLINE=96044663; PubMed=7559153;
RA Krongrad A., Wilson J.D., McPhaul M.J.;
RT "Cloning and partial sequence of the rabbit androgen receptor:
RT expression in fetal urogenital tissues.";
RL J. Androl. 16:209-212(1995).
CC -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3
CC subfamily.
CC ----------

CC ----------
DR EMBL; U16366; AAC48469.1; -.
DR HSSP; P06536; 1GDC.
DR InterPro; IPR001103; Andrgn_receptor.
DR InterPro; IPR000536; Hormone_rec_lig.
DR InterPro; IPR001628; Znf_C4steroid.
DR Pfam; PF02166; Androgen_recep; 1.
DR Pfam; PF00104; hormone_rec; 1.
DR Pfam; PF00105; zf-C4; 1.
DR PRINTS; PR00047; STROIDFINGER.
DR ProDom; PD000035; Znf_C4steroid; 1.
DR SMART; SM00430; HOLI; 1.
DR SMART; SM00399; ZnF_C4; 1.
DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW Zinc-finger; Steroid-binding.
FT NON_TER 1 1
FT DOMAIN <1 347 MODULATING.
FT DNA_BIND 349 414 NUCLEAR RECEPTOR-TYPE.

```
FT   ZN_FING      349    369       C4-TYPE.
FT   ZN_FING      385    409       C4-TYPE.
FT   DOMAIN       480    709       LIGAND-BINDING.
FT   DOMAIN         3      6       POLY-GLN.
FT   DOMAIN       182    187       POLY-PRO.
FT   DOMAIN       201    207       POLY-ALA.
FT   DOMAIN       254    262       POLY-GLY.
SQ   SEQUENCE   709 AA;  77391 MW;  40E7666137E97B6B CRC64;

  Query Match             98.4%;  Score 1349.5;  DB 1;  Length 709;
  Best Local Similarity   99.6%;  Pred. No. 2.4e-118;
  Matches  258;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        452 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 510

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 570

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        571 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 630

Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        631 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 690

Qy        242 VQVPKILSGKVKPIYFHTQ 260
              ||||||||||||||||||
Db        691 VQVPKILSGKVKPIYFHTQ 709


RESULT 2
ANDR_EULFC
ID   ANDR_EULFC     STANDARD;      PRT;   884 AA.
AC   O97776;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Androgen receptor (Dihydrotestosterone receptor).
GN   AR OR NR3C4.
OS   Eulemur fulvus collaris (Collared brown lemur).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Strepsirhini; Lemuridae; Eulemur.
OX   NCBI_TaxID=47178;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98404153; PubMed=9732460;
RA   Choong C.S., Kemppainen J.A., Wilson E.M.;
RT   "Evolution of the primate androgen receptor: a structural basis for
RT   disease.";
RL   J. Mol. Evol. 47:334-342(1998).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3
CC       subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U94178; AAC73049.1; -.
DR   HSSP; P06536; 1GDC.
DR   TRANSFAC; T04655; -.
DR   InterPro; IPR001103; Andrgn_receptor.
DR   InterPro; IPR000536; Hormone_rec_lig.
DR   InterPro; IPR001723; Stdhrmn_receptor.
DR   InterPro; IPR001628; Znf_C4steroid.
DR   Pfam; PF02166; Androgen_recep; 1.
DR   Pfam; PF00104; hormone_rec; 1.
DR   Pfam; PF00105; zf-C4; 1.
DR   PRINTS; PR00398; STRDHORMONER.
DR   PRINTS; PR00047; STROIDFINGER.
DR   ProDom; PD000035; Znf_C4steroid; 1.
DR   SMART; SM00430; HOLI; 1.
DR   SMART; SM00399; ZnF_C4; 1.
DR   PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW   Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW   Zinc-finger; Steroid-binding.
FT   DOMAIN         1    522       MODULATING (BY SIMILARITY).
FT   DNA_BIND     524    589       NUCLEAR RECEPTOR-TYPE.
FT   ZN_FING      524    544       C4-TYPE.
FT   ZN_FING      560    584       C4-TYPE.
FT   DOMAIN       655    884       LIGAND-BINDING.
FT   DOMAIN        55     58       POLY-GLN.
FT   DOMAIN        64     70       POLY-GLN.
FT   DOMAIN       116    120       POLY-ALA.
FT   DOMAIN       174    178       POLY-GLN.
FT   DOMAIN       353    362       POLY-PRO.
FT   DOMAIN       379    383       POLY-ALA.
FT   DOMAIN       408    411       POLY-ALA.
FT   DOMAIN       430    435       POLY-GLY.
SQ   SEQUENCE   884 AA;  95610 MW;  18F570E352F4D2BD CRC64;

  Query Match             98.4%;  Score 1349.5;  DB 1;  Length 884;
  Best Local Similarity   99.6%;  Pred. No. 3.1e-118;
  Matches  258;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 685

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 745

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 805

Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        806 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 865

Qy        242 VQVPKILSGKVKPIYFHTQ 260
              ||||||||||||||||||
Db        866 VQVPKILSGKVKPIYFHTQ 884


RESULT 3
ANDR_MACFA
ID   ANDR_MACFA     STANDARD;      PRT;   895 AA.
AC   O97952;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Androgen receptor (Dihydrotestosterone receptor).
GN   AR OR NR3C4.
OS   Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9541;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98404153; PubMed=9732460;
```

```
RA   Choong C.S., Kemppainen J.A., Wilson E.M.;
RT   "Evolution of the primate androgen receptor: a structural basis for
RT   disease.";
RL   J. Mol. Evol. 47:334-342(1998).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3
CC       subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U94179; AAC73050.1; -.
DR   HSSP; P06536; 1GDC.
DR   TRANSFAC; T04654; -.
DR   InterPro; IPR001103; Andrgn_receptor.
DR   InterPro; IPR000536; Hormone_rec_lig.
DR   InterPro; IPR001628; Znf_C4steroid.
DR   Pfam; PF02166; Androgen_recep; 1.
DR   Pfam; PF00104; hormone_rec; 1.
DR   Pfam; PF00105; zf-C4; 1.
DR   PRINTS; PR00047; STROIDFINGER.
DR   ProDom; PD000035; Znf_C4steroid; 1.
DR   SMART; SM00430; HOLI; 1.
DR   SMART; SM00399; ZnF_C4; 1.
DR   PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW   Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW   Zinc-finger; Steroid-binding.
FT   DOMAIN        1    533       MODULATING (BY SIMILARITY).
FT   DNA_BIND    535    600       NUCLEAR RECEPTOR-TYPE.
FT   ZN_FING     535    555       C4-TYPE.
FT   ZN_FING     571    595       C4-TYPE.
FT   DOMAIN      666    895       LIGAND-BINDING.
FT   DOMAIN       55     62       POLY-GLN.
FT   DOMAIN       68     74       POLY-GLN.
FT   DOMAIN      178    182       POLY-GLN.
FT   DOMAIN      357    366       POLY-PRO.
FT   DOMAIN      381    387       POLY-ALA.
FT   DOMAIN      434    448       POLY-GLY.
SQ   SEQUENCE   895 AA;  96494 MW;  A3EB17916F43A097 CRC64;

  Query Match             98.4%;  Score 1349.5;  DB 1;  Length 895;
  Best Local Similarity   99.6%;  Pred. No. 3.2e-118;
  Matches  258;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        638 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 696

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        697 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 756

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        757 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 816

Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        817 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 876

Qy        242 VQVPKILSGKVKPIYFHTQ 260
              ||||||||||||||||||
Db        877 VQVPKILSGKVKPIYFHTQ 895
```

```
RESULT 4
ANDR_PAPHA
ID   ANDR_PAPHA     STANDARD;      PRT;   895 AA.
AC   O97960;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Androgen receptor (Dihydrotestosterone receptor).
GN   AR OR NR3C4.
OS   Papio hamadryas (Hamadryas baboon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Papio.
OX   NCBI_TaxID=9557;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98404153; PubMed=9732460;
RA   Choong C.S., Kemppainen J.A., Wilson E.M.;
RT   "Evolution of the primate androgen receptor: a structural basis for
RT   disease.";
RL   J. Mol. Evol. 47:334-342(1998).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3
CC       subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U94176; AAC73047.1; -.
DR   HSSP; P06536; 1GDC.
DR   TRANSFAC; T04652; -.
DR   InterPro; IPR001103; Andrgn_receptor.
DR   InterPro; IPR000536; Hormone_rec_lig.
DR   InterPro; IPR001628; Znf_C4steroid.
DR   Pfam; PF02166; Androgen_recep; 1.
DR   Pfam; PF00104; hormone_rec; 1.
DR   Pfam; PF00105; zf-C4; 1.
DR   PRINTS; PR00047; STROIDFINGER.
DR   ProDom; PD000035; Znf_C4steroid; 1.
DR   SMART; SM00430; HOLI; 1.
DR   SMART; SM00399; ZnF_C4; 1.
DR   PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW   Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW   Zinc-finger; Steroid-binding.
FT   DOMAIN        1    533       MODULATING (BY SIMILARITY).
FT   DNA_BIND    535    600       NUCLEAR RECEPTOR-TYPE.
FT   ZN_FING     535    555       C4-TYPE.
FT   ZN_FING     571    595       C4-TYPE.
FT   DOMAIN      666    895       LIGAND-BINDING.
FT   DOMAIN       55     63       POLY-GLN.
FT   DOMAIN       69     74       POLY-GLN.
FT   DOMAIN      178    182       POLY-GLN.
FT   DOMAIN      357    366       POLY-PRO.
FT   DOMAIN      381    387       POLY-ALA.
FT   DOMAIN      434    448       POLY-GLY.
SQ   SEQUENCE   895 AA;  96478 MW;  9020C0DC67F11E5D CRC64;

  Query Match             98.4%;  Score 1349.5;  DB 1;  Length 895;
  Best Local Similarity   99.6%;  Pred. No. 3.2e-118;
  Matches  258;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;
```

```
Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        638 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 696

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        697 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 756

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        757 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 816

Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        817 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 876

Qy        242 VQVPKILSGKVKPIYFHTQ 260
              ||||||||||||||||||
Db        877 VQVPKILSGKVKPIYFHTQ 895


RESULT 5
ANDR_MOUSE
ID   ANDR_MOUSE     STANDARD;      PRT;   899 AA.
AC   P19091;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Androgen receptor (Dihydrotestosterone receptor).
GN   AR OR NR3C4.
OS   Mus musculus (Mouse).
QC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c;
RX   MEDLINE=90386642; PubMed=2403358;
RA   He W.W., Fischer L.M., Sun S., Bilhartz D.L., Zhu X., Young C.Y.F.,
RA   Kelley D.B., Tindall D.J.;
RT   "Molecular cloning of androgen receptors from divergent species with
RT   a polymerase chain reaction technique: complete cDNA sequence of the
RT   mouse androgen receptor and isolation of androgen receptor cDNA
RT   probes from dog, guinea pig and clawed frog.";
RL   Biochem. Biophys. Res. Commun. 171:697-704(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91133433; PubMed=2178222;
RA   Gaspar M.L., Meo T., Tosi M.;
RT   "Structure and size distribution of the androgen receptor mRNA in
RT   wild-type and Tfm/Y mutant mice.";
RL   Mol. Endocrinol. 4:1600-1610(1990).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91354214; PubMed=1883336;
RA   Faber P.W., King A., van Rooij H.C.J., Brinkmann A.O., de Both N.J.,
RA   Trapman J.;
RT   "The mouse androgen receptor. Functional analysis of the protein and
RT   characterization of the gene.";
RL   Biochem. J. 278:269-278(1991).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92017874; PubMed=1681426;
RA   Charest N.J., Zhou Z., Lubahn D.B., Olsen K.L., Wilson E.M.,
RA   French F.S.;
RT   "A frameshift mutation destabilizes androgen receptor messenger RNA
RT   in the Tfm mouse.";
RL   Mol. Endocrinol. 5:573-581(1991).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- MISCELLANEOUS: IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS
CC       ARE THOUGHT TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS;
CC       HORMONE BINDING GREATLY INCREASES RECEPTOR AFFINITY. THE
CC       HORMONE-RECEPTOR COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA
CC       SEQUENCES UPSTREAM OF TRANSCRIPTIONAL START SITES.
CC   -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3
CC       subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; S56585; AAB19916.1; -.
DR   EMBL; X53779; CAA37795.1; -.
DR   EMBL; M37890; AAA37234.1; -.
DR   EMBL; X59592; CAA42160.1; -.
DR   PIR; A35895; A35895.
DR   HSSP; P06536; 1GDC.
DR   TRANSFAC; T00041; -.
DR   MGD; MGI:88064; Ar.
DR   InterPro; IPR001103; Andrgn_receptor.
DR   InterPro; IPR000536; Hormone_rec_lig.
DR   InterPro; IPR001628; Znf_C4steroid.
DR   Pfam; PF02166; Androgen_recep; 1.
DR   Pfam; PF00104; hormone_rec; 1.
DR   Pfam; PF00105; zf-C4; 1.
DR   PRINTS; PR00047; STROIDFINGER.
DR   ProDom; PD000035; Znf_C4steroid; 1.
DR   SMART; SM00430; HOLI; 1.
DR   SMART; SM00399; ZnF_C4; 1.
DR   PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW   Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW   Zinc-finger; Steroid-binding.
FT   DOMAIN       1    537       MODULATING (BY SIMILARITY).
FT   DNA_BIND   539    604       NUCLEAR RECEPTOR-TYPE.
FT   ZN_FING    539    559       C4-TYPE.
FT   ZN_FING    575    599       C4-TYPE.
FT   DOMAIN     670    899       LIGAND-BINDING.
FT   DOMAIN      63     67       POLY-ARG.
FT   DOMAIN     174    193       POLY-GLN.
FT   DOMAIN     367    373       POLY-PRO.
FT   DOMAIN     391    397       POLY-ALA.
FT   DOMAIN     441    447       POLY-GLY.
SQ   SEQUENCE   899 AA;  98193 MW;  FD9EE07C07F7A568 CRC64;

  Query Match            98.4%;  Score 1349.5;  DB 1;  Length 899;
  Best Local Similarity  99.6%;  Pred. No. 3.2e-118;
  Matches  258;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        642 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 700

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        701 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 760

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        761 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 820

Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        821 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 880

Qy        242 VQVPKILSGKVKPIYFHTQ 260
              ||||||||||||||||||
Db        881 VQVPKILSGKVKPIYFHTQ 899
```

```
RESULT 6
ANDR_RAT
ID   ANDR_RAT       STANDARD;      PRT;   902 AA.
AC   P15207; Q63049;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Androgen receptor (Dihydrotestosterone receptor).
GN   AR OR NR3C4.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89112209; PubMed=3216867;
RA   Tan J., Joseph D.R., Quarmby V.E., Lubahn D.B., Sar M., French F.S.,
RA   Wilson E.M.;
RT   "The rat androgen receptor: primary structure, autoregulation of its
RT   messenger ribonucleic acid, and immunocytochemical localization of
RT   the receptor protein.";
RL   Mol. Endocrinol. 2:1276-1285(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Prostate;
RX   MEDLINE=89017168; PubMed=3174628;
RA   Chang C., Kokontis J., Liao S.;
RT   "Structural analysis of complementary DNA and amino acid sequences of
RT   human and rat androgen receptors.";
RL   Proc. Natl. Acad. Sci. U.S.A. 85:7211-7215(1988).
RN   [3]
RP   SEQUENCE FROM N.A., AND VARIANT TFM GLN-735.
RX   MEDLINE=90256822; PubMed=2341409;
RA   Yarbrough W.G., Quarmby V.E., Simental J.A., Joseph D.R., Sar M.,
RA   Lubahn D.B., Olsen K.L., French F.S., Wilson E.M.;
RT   "A single base mutation in the androgen receptor gene causes androgen
RT   insensitivity in the testicular feminized rat.";
RL   J. Biol. Chem. 265:8893-8900(1990).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- TISSUE SPECIFICITY: HIGHEST LEVELS IN THE SEMINAL VESICLE,
CC       VENTRAL PROSTATE AND COAGULATING GLAND WITH LOWER LEVELS IN THE
CC       KIDNEY, AND LEVATOR ANI MUSCLE.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- DISEASE: Defects in AR are a cause of androgen insensitivity. Rats
CC       with this syndrome are called testicular feminized (TFM).
CC   -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3
CC       subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M20133; AAA40733.1; -.
DR   EMBL; M23264; AAA40759.1; -.
DR   EMBL; J05454; AAA40734.1; -.
DR   PIR; B40494; B40494.
DR   PDB; 1I37; 02-MAY-01.
DR   PDB; 1I38; 02-MAY-01.
DR   TRANSFAC; T00042; -.
DR   InterPro; IPR001103; Andrgn_receptor.
DR   InterPro; IPR000536; Hormone_rec_lig.
DR   InterPro; IPR001628; Znf_C4steroid.
DR   Pfam; PF02166; Androgen_recep; 1.
DR   Pfam; PF00104; hormone_rec; 1.
DR   Pfam; PF00105; zf-C4; 1.
DR   PRINTS; PR00047; STROIDFINGER.
DR   ProDom; PD000035; Znf_C4steroid; 1.
DR   SMART; SM00430; HOLI; 1.
DR   SMART; SM00399; ZnF_C4; 1.
DR   PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW   Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW   Zinc-finger; Steroid-binding; Disease mutation; 3D-structure.
FT   DOMAIN        1    540       MODULATING.
FT   DNA_BIND    542    607       NUCLEAR RECEPTOR-TYPE.
FT   ZN_FING     542    562       C4-TYPE.
FT   ZN_FING     578    602       C4-TYPE.
FT   DOMAIN      673    902       LIGAND-BINDING.
FT   DOMAIN       63     67       POLY-ARG.
FT   DOMAIN      174    195       POLY-GLN.
FT   DOMAIN      370    376       POLY-PRO.
FT   DOMAIN      394    400       POLY-ALA.
FT   DOMAIN      444    450       POLY-GLY.
FT   VARIANT     735    735       R -> Q (IN TFM; HAS ONLY 10-15% OF THE
FT                                ANDROGEN-BINDING CAPACITY OF WILD-TYPE
FT                                AR).
FT   CONFLICT    195    195       MISSING (IN REF. 3).
FT   CONFLICT    389    389       S -> L (IN REF. 2).
SQ   SEQUENCE   902 AA;  98217 MW;  43F4064759FDCBED CRC64;

  Query Match             98.4%;  Score 1349.5;  DB 1;  Length 902;
  Best Local Similarity   99.6%;  Pred. No. 3.2e-118;
  Matches  258;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        645 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 703

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        704 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 763

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        764 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 823

Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        824 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 883

Qy        242 VQVPKILSGKVKPIYFHTQ 260
              |||||||||||||||||||
Db        884 VQVPKILSGKVKPIYFHTQ 902


RESULT 7
ANDR_CANFA
ID   ANDR_CANFA     STANDARD;      PRT;   907 AA.
AC   Q9TT90;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Androgen receptor (Dihydrotestosterone receptor).
GN   AR OR NR3C4.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21618348; PubMed=11768233;
RA   Lu B., Smock S.L., Castleberry T.A., Owen T.A.;
RT   "Molecular cloning and functional characterization of the canine
RT   androgen receptor.";
RL   Mol. Cell. Biochem. 226:129-140(2001).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
```

```
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3
CC       subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF197950; AAF18084.1; -.
DR   HSSP; P06536; 1GDC.
DR   InterPro; IPR001103; Andrgn_receptor.
DR   InterPro; IPR000536; Hormone_rec_lig.
DR   InterPro; IPR001628; Znf_C4steroid.
DR   Pfam; PF02166; Androgen_recep; 1.
DR   Pfam; PF00104; hormone_rec; 1.
DR   Pfam; PF00105; zf-C4; 1.
DR   PRINTS; PR00047; STROIDFINGER.
DR   ProDom; PD000035; Znf_C4steroid; 1.
DR   SMART; SM00430; HOLI; 1.
DR   SMART; SM00399; ZnF_C4; 1.
DR   PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW   Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW   Zinc-finger; Steroid-binding.
FT   DOMAIN        1    545       MODULATING (BY SIMILARITY).
FT   DNA_BIND    547    612       NUCLEAR RECEPTOR-TYPE.
FT   ZN_FING     547    567       C4-TYPE.
FT   ZN_FING     583    607       C4-TYPE.
FT   DOMAIN      678    907       LIGAND-BINDING.
FT   DOMAIN       55     64       POLY-GLN.
FT   DOMAIN       70     76       POLY-GLN.
FT   DOMAIN      131    134       POLY-GLN.
FT   DOMAIN      180    202       POLY-GLN.
FT   DOMAIN      329    332       POLY-SER.
FT   DOMAIN      375    384       POLY-PRO.
FT   DOMAIN      399    405       POLY-ALA.
SQ   SEQUENCE   907 AA;  98726 MW;  C8619F78DD2338AF CRC64;

  Query Match           98.4%;  Score 1349.5;  DB 1;  Length 907;
  Best Local Similarity   99.6%;  Pred. No. 3.2e-118;
  Matches  258;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        650 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 708

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        709 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 768

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        769 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 828

Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        829 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 888

Qy        242 VQVPKILSGKVKPIYFHTQ 260
              ||||||||||||||||||
Db        889 VQVPKILSGKVKPIYFHTQ 907


RESULT 8
ANDR_PANTR
ID   ANDR_PANTR     STANDARD;      PRT;   911 AA.
AC   O97775;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Androgen receptor (Dihydrotestosterone receptor).
GN   AR OR NR3C4.
OS   Pan troglodytes (Chimpanzee).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX   NCBI_TaxID=9598;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98404153; PubMed=9732460;
RA   Choong C.S., Kemppainen J.A., Wilson E.M.;
RT   "Evolution of the primate androgen receptor: a structural basis for
RT   disease.";
RL   J. Mol. Evol. 47:334-342(1998).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3
CC       subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U94177; AAC73048.1; -.
DR   HSSP; P06536; 1GDC.
DR   TRANSFAC; T04653; -.
DR   InterPro; IPR001103; Andrgn_receptor.
DR   InterPro; IPR000536; Hormone_rec_lig.
DR   InterPro; IPR001628; Znf_C4steroid.
DR   Pfam; PF02166; Androgen_recep; 1.
DR   Pfam; PF00104; hormone_rec; 1.
DR   Pfam; PF00105; zf-C4; 1.
DR   PRINTS; PR00047; STROIDFINGER.
DR   ProDom; PD000035; Znf_C4steroid; 1.
DR   SMART; SM00430; HOLI; 1.
DR   SMART; SM00399; ZnF_C4; 1.
DR   PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW   Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW   Zinc-finger; Steroid-binding.
FT   DOMAIN        1    549       MODULATING (BY SIMILARITY).
FT   DNA_BIND    551    616       NUCLEAR RECEPTOR-TYPE.
FT   ZN_FING     551    571       C4-TYPE.
FT   ZN_FING     587    611       C4-TYPE.
FT   DOMAIN      682    911       LIGAND-BINDING.
FT   DOMAIN       57     78       POLY-GLN.
FT   DOMAIN       84     88       POLY-GLN.
FT   DOMAIN      192    196       POLY-GLN.
FT   DOMAIN      371    380       POLY-PRO.
FT   DOMAIN      395    401       POLY-ALA.
FT   DOMAIN      448    464       POLY-GLY.
SQ   SEQUENCE   911 AA;  98402 MW;  601B9BD4E697DAA4 CRC64;

  Query Match           98.4%;  Score 1349.5;  DB 1;  Length 911;
  Best Local Similarity   99.6%;  Pred. No. 3.2e-118;
  Matches  258;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        654 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 712

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 772
```

```
Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        773 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 832

Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        833 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 892

Qy        242 VQVPKILSGKVKPIYFHTQ 260
              ||||||||||||||||||
Db        893 VQVPKILSGKVKPIYFHTQ 911


RESULT 9
ANDR_HUMAN
ID   ANDR_HUMAN     STANDARD;      PRT;   919 AA.
AC   P10275;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Androgen receptor (Dihydrotestosterone receptor).
GN   AR OR NR3C4 OR DHTR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89112208; PubMed=3216866;
RA   Lubahn D.B., Joseph D.R., Sar M., Tan J., Higgs H.N., Larson R.E.,
RA   French F.S., Wilson E.M.;
RT   "The human androgen receptor: complementary deoxyribonucleic acid
RT   cloning, sequence analysis and gene expression in prostate.";
RL   Mol. Endocrinol. 2:1265-1275(1988).
RN   [2]
RP   SEQUENCE FROM N.A., AND VARIANT CAIS MET-866.
RX   MEDLINE=90083302; PubMed=2594783;
RA   Lubahn D.B., Brown T.R., Simental J.A., Higgs H.N., Migeon C.J.,
RA   Wilson E.M., French F.S.;
RT   "Sequence of the intron/exon junctions of the coding region of the
RT   human androgen receptor gene and identification of a point mutation
RT   in a family with complete androgen insensitivity.";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:9534-9538(1989).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90258935; PubMed=2342476;
RA   Govindan M.V.;
RT   "Specific region in hormone binding domain is essential for hormone
RT   binding and trans-activation by human androgen receptor.";
RL   Mol. Endocrinol. 4:417-427(1990).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Prostate;
RX   MEDLINE=89017168; PubMed=3174628;
RA   Chang C., Kokontis J., Liao S.;
RT   "Structural analysis of complementary DNA and amino acid sequences of
RT   human and rat androgen receptors.";
RL   Proc. Natl. Acad. Sci. U.S.A. 85:7211-7215(1988).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Prostate;
RX   MEDLINE=89098909; PubMed=2911578;
RA   Tilley W.D., Marcelli M., Wilson J.D., McPhaul M.J.;
RT   "Characterization and expression of a cDNA encoding the human androgen
RT   receptor.";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:327-331(1989).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Prostate;
RX   MEDLINE=91155943; PubMed=2293020;
RA   Marcelli M., Tilley W.D., Wilson C.M., Griffin J.E., Wilson J.D.,
RA   McPhaul M.J.;
RT   "Definition of the human androgen receptor gene structure permits the
RT   identification of mutations that cause androgen resistance: premature
RT   termination of the receptor protein at amino acid residue 588 causes
RT   complete androgen resistance.";
RL   Mol. Endocrinol. 4:1105-1116(1990).
RN   [7]
RP   SEQUENCE OF 189-919 FROM N.A.
RX   MEDLINE=88178111; PubMed=3353726;
RA   Chang C., Kokontis J., Liao S.;
RT   "Molecular cloning of human and rat complementary DNA encoding
RT   androgen receptors.";
RL   Science 240:324-326(1988).
RN   [8]
RP   SEQUENCE OF 468-919 FROM N.A.
RX   MEDLINE=88240407; PubMed=3377788;
RA   Trapman J., Klaassen P., Kuiper G.G.J.M., van der Korput J.A.G.M.,
RA   Faber P.W., van Rooij H.C.J., Geurts van Kessel A., Voorhorst M.M.,
RA   Mulder E., Brinkmann A.O.;
RT   "Cloning, structure and expression of a cDNA encoding the human
RT   androgen receptor.";
RL   Biochem. Biophys. Res. Commun. 153:241-248(1988).
RN   [9]
RP   INTERACTION WITH RAN.
RX   MEDLINE=99329028; PubMed=10400640;
RA   Hsiao P.-W., Lin D.-L., Nakao R., Chang C.;
RT   "The linkage of Kennedy's neuron disease to ARA24, the first
RT   identified androgen receptor polyglutamine region-associated
RT   coactivator.";
RL   J. Biol. Chem. 274:20229-20234(1999).
RN   [10]
RP   POLYMORPHISM OF POLY-GLN REGION.
RX   MEDLINE=92220629; PubMed=1561105;
RA   Sleddens H.F.B.M., Oostra B.A., Brinkmann A.O., Trapman J.;
RT   "Trinucleotide repeat polymorphism in the androgen receptor gene
RT   (AR).";
RL   Nucleic Acids Res. 20:1427-1427(1992).
RN   [11]
RP   POLYMORPHISM OF POLY-GLY REGION.
RC   TISSUE=Blood;
RA   Lu J., Danielsen M.;
RL   Submitted (FEB-1995) to the EMBL/GenBank/DDBJ databases.
RN   [12]
RP   POLYMORPHISM OF POLY-GLN REGION.
RX   MEDLINE=97250535; PubMed=9096391;
RA   Giovannucci E., Stampfer M.J., Krithivas K., Brown M., Brufsky A.,
RA   Talcott J., Hennekens C.H., Kantoff P.W.;
RT   "The CAG repeat within the androgen receptor gene and its
RT   relationship to prostate cancer.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:3320-3323(1997).
RN   [13]
RP   ERRATUM.
RA   Giovannucci E., Stampfer M.J., Krithivas K., Brown M., Dahl D.,
RA   Brufsky A., Talcott J., Hennekens C.H., Kantoff P.W.;
RL   Proc. Natl. Acad. Sci. U.S.A. 94:8272-8272(1997).
RN   [14]
RP   REVIEW ON VARIANTS.
RX   MEDLINE=93092459; PubMed=1458719;
RA   Pinsky L., Trifiro M.A., Kaufman M., Beitel L.K., Mhatre A.,
RA   Kazemi-Esfarjani P., Sabbaghian N., Lumbroso R., Alvarado C.,
RA   Vasiliou M., Gottlieb B.;
RT   "Androgen resistance due to mutation of the androgen receptor.";
RL   Clin. Invest. Med. 15:456-472(1992).
RN   [15]
RP   REVIEW ON VARIANTS AIS.
RX   MEDLINE=93339360; PubMed=8339746;
RA   Brown T.R., Scherer P.A., Chang Y.-T., Migeon C.J., Ghirri P.,
RA   Murono K., Zhou Z.;
RT   "Molecular genetics of human androgen insensitivity.";
RL   Eur. J. Pediatr. 152 Suppl. 2:S62-S69(1993).
RN   [16]
RP   REVIEW ON VARIANTS.
RX   MEDLINE=94059770; PubMed=8240973;
RA   Sultan C., Lumbroso S., Poujol N., Belon C., Boudon C.,
```

```
RA   Lobaccaro J.-M.;
RT   "Mutations of androgen receptor gene in androgen insensitivity
RT   syndromes.";
RL   J. Steroid Biochem. Mol. Biol. 46:519-530(1993).
RN   [17]
RP   REVIEW ON VARIANTS.
RX   MEDLINE=95023089; PubMed=7937057;
RA   Patterson M.N., Hughes I.A., Gottlieb B., Pinsky L.;
RT   "The androgen receptor gene mutations database.";
RL   Nucleic Acids Res. 22:3560-3562(1994).
RN   [18]
RP   REVIEW ON VARIANTS.
RX   MEDLINE=95352489; PubMed=7626493;
RA   Brinkmann A.O., Jenster G., Ris-Stalpers C., van der Korput J.A.G.M.,
RA   Bruggenwirth H.T., Boehmer A.L.M., Trapman J.;
RT   "Androgen receptor mutations.";
RL   J. Steroid Biochem. Mol. Biol. 53:443-448(1995).
RN   [19]
RP   REVIEW ON VARIANTS.
RX   MEDLINE=97169385; PubMed=9016528;
RA   Gottlieb B., Trifiro M.A., Lumbroso R., Vasiliou D.M., Pinsky L.;
RT   "The androgen receptor gene mutations database.";
RL   Nucleic Acids Res. 25:158-162(1997).
RN   [20]
RP   VARIANT LNCAP ALA-877.
RX   MEDLINE=91083633; PubMed=2260966;
RA   Veldscholte J., Ris-Stalpers C., Kuiper G.G.J.M., Jenster G.,
RA   Berrevoets C.A., Claassen E., van Rooij H.C.J., Trapman J.,
RA   Brinkmann A.O., Mulder E.;
RT   "A mutation in the ligand binding domain of the androgen receptor of
RT   human LNCaP cells affects steroid binding characteristics and
RT   response to anti-androgens.";
RL   Biochem. Biophys. Res. Commun. 173:534-540(1990).
RN   [21]
RP   VARIANTS CAIS CYS-774; GLN-831 AND MET-866.
RX   MEDLINE=91186983; PubMed=2082179;
RA   Brown T.R., Lubahn D.B., Wilson E.M., French F.S., Migeon C.J.,
RA   Corfen J.L.;
RT   "Functional characterization of naturally occurring mutant androgen
RT   receptors from subjects with complete androgen insensitivity.";
RL   Mol. Endocrinol. 4:1759-1772(1990).
RN   [22]
RP   VARIANT CYS-774.
RX   MEDLINE=91310758; PubMed=1856263;
RA   Marcelli M., Tilley W.D., Zoppi S., Griffin J.E., Wilson J.D.,
RA   McPhaul M.J.;
RT   "Androgen resistance associated with a mutation of the androgen
RT   receptor at amino acid 772 (Arg-->Cys) results from a combination of
RT   decreased messenger ribonucleic acid levels and impairment of
RT   receptor function.";
RL   J. Clin. Endocrinol. Metab. 73:318-325(1991).
RN   [23]
RP   VARIANT CAIS PRO-617.
RX   MEDLINE=91154385; PubMed=1999491;
RA   Marcelli M., Zoppi S., Grino P.B., Griffin J.E., Wilson J.D.,
RA   McPhaul M.J.;
RT   "A mutation in the DNA-binding domain of the androgen receptor gene
RT   causes complete testicular feminization in a patient with
RT   receptor-positive androgen resistance.";
RL   J. Clin. Invest. 87:1123-1126(1991).
RN   [24]
RP   VARIANT PAIS CYS-763.
RX   MEDLINE=91185626; PubMed=2010552;
RA   McPhaul M.J., Marcelli M., Tilley W.D., Griffin J.E.,
RA   Isidro-Gutierrez R.F., Wilson J.D.;
RT   "Molecular basis of androgen resistance in a family with a qualitative
RT   abnormality of the androgen receptor and responsive to high-dose
RT   androgen therapy.";
RL   J. Clin. Invest. 87:1413-1421(1991).
RN   [25]

  Query Match             98.4%;  Score 1349.5;  DB 1;  Length 919;
  Best Local Similarity   99.6%;  Pred. No. 3.3e-118;

  Matches  258;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 720

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 780

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 840

Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 900

Qy        242 VQVPKILSGKVKPIYFHTQ 260
              |||||||||||||||||||
Db        901 VQVPKILSGKVKPIYFHTQ 919


RESULT 10
PRGR_RABIT
ID   PRGR_RABIT     STANDARD;      PRT;   930 AA.
AC   P06186;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-JAN-1988 (Rel. 06, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Progesterone receptor (PR).
GN   PGR OR NR3C3.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87067449; PubMed=3538016;
RA   Loosfelt H., Atger M., Misrahi M., Guiochon-Mantel A., Meriel C.,
RA   Logeat F., Benarous R., Milgrom E.;
RT   "Cloning and sequence analysis of rabbit progesterone-receptor
RT   complementary DNA.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:9045-9049(1986).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3
CC       subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M14547; AAA31443.1; -.
DR   PIR; A25923; A25923.
DR   HSSP; P06401; 1A28.
DR   TRANSFAC; T00697; -.
DR   InterPro; IPR000536; Hormone_rec_lig.
DR   InterPro; IPR000128; Progest_receptor.
DR   InterPro; IPR001723; Stdhrmn_receptor.
DR   InterPro; IPR001628; Znf_C4steroid.
DR   Pfam; PF00104; hormone_rec; 1.
DR   Pfam; PF02161; Prog_receptor; 1.
DR   Pfam; PF00105; zf-C4; 1.
DR   PRINTS; PR00398; STRDHORMONER.
```

```
DR   PRINTS; PR00047; STROIDFINGER.
DR   ProDom; PD000035; Znf_C4steroid; 1.
DR   SMART; SM00430; HOLI; 1.
DR   SMART; SM00399; ZnF_C4; 1.
DR   PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW   Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW   Zinc-finger; Steroid-binding.
FT   DOMAIN        1    565       MODULATING, PRO-RICH.
FT   DNA_BIND    568    633       NUCLEAR RECEPTOR-TYPE.
FT   ZN_FING     568    588       C4-TYPE.
FT   ZN_FING     604    628       C4-TYPE.
FT   DOMAIN      678    930       STEROID-BINDING.
FT   DOMAIN      184    188       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ   SEQUENCE   930 AA;  98666 MW;  644FF4C13BF2F883 CRC64;

  Query Match             55.9%;  Score 766;  DB 1;  Length 930;
  Best Local Similarity   55.5%;  Pred. No. 1.2e-63;
  Matches  137;  Conservative   57;  Mismatches   53;  Indels    0;  Gaps    0;

Qy         12 PIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFRNLHV 71
              |  :|:| :||| |: ||||| :|| :: ::||:|||:|||||: ||||:|:|||||||:
Db        682 PPLINLLMSIEPDVIYAGHDNTKPDTSSSLLTSLNQLGERQLLSVVKWSKSLPGFRNLHI 741

Qy         72 DDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHL 131
              |||: :|||||| |||| :|||| :|: :|||||||: || || :|  || |: |  :
Db        742 DDQITLIQYSWMSLMVFGLGWRSYKHVSGQMLYFAPDLILNEQRMKESSFYSLCLTMWQI 801

Qy        132 SQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPT 191
              |||  ||:: :|||||| ||| : ||::||::|  |:|:| :||:|| : |  ::|
Db        802 PQEFVKLQVSQEEFLCMKVLLLLNTIPLEGLRSQSQFEEMRSSYIRELIKAIGLRQKGVV 861

Qy        192 SCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQVPKILSGK 251
              | |:||||||||||::  : ::|| :  :  |:|  :||:|||||:|:|: |:||||:|
Db        862 SSSQRFYQLTKLLDNLHDLVKQLHLYCLNTFIQSRALSVEFPEMMSEVIAAQLPKILAGM 921

Qy        252 VKPIYFH 258
              |||: ||
Db        922 VKPLLFH 928


RESULT 11
PRGR_HUMAN
ID   PRGR_HUMAN     STANDARD;      PRT;   933 AA.
AC   P06401; Q9UPF7;
DT   01-JAN-1988 (Rel. 06, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Progesterone receptor (PR).
GN   PGR OR NR3C3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX   MEDLINE=90228361; PubMed=2328727;
RA   Kastner P., Krust A., Turcotte B., Stropp U., Tora L., Gronemeyer H.,
RA   Chambon P.;
RT   "Two distinct estrogen-regulated promoters generate transcripts
RT   encoding the two functionally different human progesterone receptor
RT   forms A and B.";
RL   EMBO J. 9:1603-1614(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87184565; PubMed=3551956;
RA   Misrahi M., Atger M., D'Auriol L., Loosfelt H., Meriel C.,
RA   Fridlansky F., Guiochon-Mantel A., Galibert F., Milgrom E.;
RT   "Complete amino acid sequence of the human progesterone receptor
RT   deduced from cloned cDNA.";
RL   Biochem. Biophys. Res. Commun. 143:740-748(1987).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Kieback D.G., Agoulnik I.U., Tong X.-W.;
RL   Submitted (JUL-1997) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 682-933.
RX   MEDLINE=98282128; PubMed=9620806;
RA   Williams S.P., Sigler P.B.;
RT   "Atomic structure of progesterone complexed with its receptor.";
RL   Nature 393:392-396(1998).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=B;
CC         IsoId=P06401-1; Sequence=Displayed;
CC       Name=A;
CC         IsoId=P06401-2; Sequence=VSP_003706;
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3
CC       subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X51730; CAA36018.1; -.
DR   EMBL; M15716; AAA60081.1; -.
DR   EMBL; AF016381; AAD01587.1; -.
DR   PIR; S09971; QRHUP.
DR   PDB; 1A28; 15-JUL-98.
DR   PDB; 1E3K; 14-JUN-01.
DR   TRANSFAC; T00696; -.
DR   Genew; HGNC:8910; PGR.
DR   MIM; 607311; -.
DR   MIM; 264080; -.
DR   GO; GO:0003707; F:steroid hormone receptor activity; TAS.
DR   GO; GO:0007267; P:cell-cell signaling; TAS.
DR   GO; GO:0007165; P:signal transduction; TAS.
DR   InterPro; IPR000536; Hormone_rec_lig.
DR   InterPro; IPR000128; Progest_receptor.
DR   InterPro; IPR001723; Stdhrmn_receptor.
DR   InterPro; IPR001628; Znf_C4steroid.
DR   Pfam; PF00104; hormone_rec; 1.
DR   Pfam; PF02161; Prog_receptor; 1.
DR   Pfam; PF00105; zf-C4; 1.
DR   PRINTS; PR00398; STRDHORMONER.
DR   PRINTS; PR00047; STROIDFINGER.
DR   ProDom; PD000035; Znf_C4steroid; 1.
DR   SMART; SM00430; HOLI; 1.
DR   SMART; SM00399; ZnF_C4; 1.
DR   PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW   Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW   Zinc-finger; Steroid-binding; Phosphorylation; Alternative splicing;
KW   3D-structure; Polymorphism.
FT   DOMAIN        1    566       MODULATING, PRO-RICH.
FT   DNA_BIND    567    632       NUCLEAR RECEPTOR-TYPE.
FT   ZN_FING     567    587       C4-TYPE.
FT   ZN_FING     603    627       C4-TYPE.
FT   DOMAIN      681    933       STEROID-BINDING.
FT   DOMAIN      183    187       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   MOD_RES      41     41       PHOSPHORYLATION (BY CK).
FT   MOD_RES     227    227       PHOSPHORYLATION (BY CK).
FT   MOD_RES     232    232       PHOSPHORYLATION (BY CK).
FT   MOD_RES     552    552       PHOSPHORYLATION (BY CK).
FT   MOD_RES     793    793       PHOSPHORYLATION (BY CK).
FT   VARSPLIC      1    164       Missing (in isoform A).
FT                                /FTId=VSP_003706.
FT   VARIANT     625    625       R -> I (IN dbSNP:2020874).
```

```
FT                                     /FTId=VAR_014627.
FT   VARIANT     865    865       S -> L (IN dbSNP:2020880).
FT                                     /FTId=VAR_014628.
FT   CONFLICT    226    226       G -> S (IN REF. 1).
FT   CONFLICT    256    256       V -> S (IN REF. 1).
FT   CONFLICT    344    344       S -> T (IN REF. 1).
FT   CONFLICT    660    660       L -> V (IN REF. 2).
FT   HELIX       686    693
FT   TURN        694    694
FT   HELIX       711    735
FT   TURN        737    738
FT   HELIX       739    741
FT   HELIX       744    770
FT   TURN        772    773
FT   STRAND      775    779
FT   TURN        780    781
FT   STRAND      782    785
FT   HELIX       786    790
FT   TURN        791    791
FT   HELIX       795    811
FT   TURN        812    812
FT   HELIX       815    826
FT   TURN        827    827
FT   STRAND      829    830
FT   TURN        832    833
FT   TURN        836    837
FT   HELIX       838    857
FT   TURN        858    859
FT   HELIX       863    896
FT   TURN        897    897
FT   HELIX       898    901
FT   TURN        902    902
FT   HELIX       907    921
FT   TURN        922    923
FT   STRAND      926    927
SQ   SEQUENCE   933 AA;  98994 MW;  80414B7F1F317F8E CRC64;

  Query Match             55.8%;  Score 765;  DB 1;  Length 933;
  Best Local Similarity   55.5%;  Pred. No. 1.5e-63;
  Matches  137;  Conservative   57;  Mismatches   53;  Indels    0;  Gaps    0;

Qy         12 PIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFRNLHV 71
              |  :|:| :||| |: ||||| :||: ::||:|||:|||||: ||||:|:|||||||:
Db        685 PPLINLLMSIEPDVIYAGHDNTKPDTSSSLLTSLNQLGERQLLSVVKWSKSLPGFRNLHI 744

Qy         72 DDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHL 131
              |||: :|||||| |||| :||||: :|: :||||||||: || || :|  || |: |  :
Db        745 DDQITLIQYSWMSLMVFGLGWRSYKHVSGQMLYFAPDLILNEQRMKESSFYSLCLTMWQI 804

Qy        132 SQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPT 191
               |||  ||:: :|||||| ||| : ||::||::|  |:|:| :||:|| : |  ::|
Db        805 PQEFVKLQVSQEEFLCMKVLLLLNTIPLEGLRSQTQFEEMRSSYIRELIKAIGLRQKGVV 864

Qy        192 SCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQVPKILSGK 251
              | |:||||||||||::  : ::|| :  :  |:|  :||:|||||:|:|: |:||||:|
Db        865 SSSQRFYQLTKLLDNLHDLVKQLHLYCLNTFIQSRALSVEFPEMMSEVIAAQLPKILAGM 924

Qy        252 VKPIYFH 258
              |||: ||
Db        925 VKPLLFH 931


RESULT 12
PRGR_RAT
ID   PRGR_RAT       STANDARD;      PRT;   923 AA.
AC   Q63449;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Progesterone receptor (PR).
GN   PGR OR NR3C3.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Placenta;
RX   MEDLINE=94130817; PubMed=8299566;
RA   Park-Sarge O.K., Mayo K.E.;
RT   "Regulation of the progesterone receptor gene by gonadotropins and
RT   cyclic adenosine 3',5'-monophosphate in rat granulosa cells.";
RL   Endocrinology 134:709-718(1994).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3
CC       subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L16922; AAA19916.1; -.
DR   PIR; I53280; I53280.
DR   HSSP; P06401; 1A28.
DR   TRANSFAC; T04681; -.
DR   InterPro; IPR000536; Hormone_rec_lig.
DR   InterPro; IPR000128; Progest_receptor.
DR   InterPro; IPR001723; Stdhrmn_receptor.
DR   InterPro; IPR001628; Znf_C4steroid.
DR   Pfam; PF00104; hormone_rec; 1.
DR   Pfam; PF02161; Prog_receptor; 1.
DR   Pfam; PF00105; zf-C4; 1.
DR   PRINTS; PR00398; STRDHORMONER.
DR   PRINTS; PR00047; STROIDFINGER.
DR   ProDom; PD000035; Znf_C4steroid; 1.
DR   SMART; SM00430; HOLI; 1.
DR   SMART; SM00399; ZnF_C4; 1.
DR   PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW   Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW   Zinc-finger; Steroid-binding.
FT   DOMAIN        1    556       MODULATING, PRO-RICH.
FT   DNA_BIND    557    622       NUCLEAR RECEPTOR-TYPE.
FT   ZN_FING     557    577       C4-TYPE.
FT   ZN_FING     593    617       C4-TYPE.
FT   DOMAIN      671    923       STEROID-BINDING.
FT   DOMAIN      184    188       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ   SEQUENCE   923 AA;  99407 MW;  05384B9656BF22DC CRC64;

  Query Match             55.4%;  Score 759;  DB 1;  Length 923;
  Best Local Similarity   55.5%;  Pred. No. 5.5e-63;
  Matches  137;  Conservative   55;  Mismatches   55;  Indels    0;  Gaps    0;

Qy         12 PIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFRNLHV 71
              |  :|:| :||| || ||||| :||: ::||:|||:|||||: ||||:|:|||||||:
Db        675 PPLINLLMSIEPDVVYAGHDNTKPDTSSSLLTSLNQLGERQLLSVVKWSKSLPGFRNLHI 734

Qy         72 DDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHL 131
              |||: :|||||| |||| :||||: :|: :||||||||: || || :   || |: |  :
Db        735 DDQITLIQYSWMSLMVFGLGWRSYKHVSGQMLYFAPDLILNEQRMKELSFYSLCLTMWQI 794

Qy        132 SQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPT 191
               |||  ||:| :|||||| ||| : ||::||::|  |:|:| :||:|| : |  ::|
Db        795 PQEFVKLQVTHEEFLCMKVLLLLNTIPLEGLRSQSQFEEMRSSYIRELIKAIGLRQKGVV 854

Qy        192 SCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQVPKILSGK 251
              |:|||||||||||:  : ::|| :  :  |:|  ::|:|||||:|:|: |:||||:|
Db        855 PSSQRFYQLTKLLDSLHDLVKQLHLYCLNTFIQSRALAVEFPEMMSEVIAAQLPKILAGM 914
```

```
Qy        252 VKPIYFH 258
              |||: ||
Db        915 VKPLLFH 921


RESULT 13
PRGR_MOUSE
ID   PRGR_MOUSE     STANDARD;      PRT;   923 AA.
AC   Q00175;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Progesterone receptor (PR).
GN   PGR OR NR3C3 OR PR.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91299759; PubMed=2069958;
RA   Schott D.R., Shyamala G., Schneider W., Parry G.;
RT   "Molecular cloning, sequence analyses, and expression of
RT   complementary DNA encoding murine progesterone receptor.";
RL   Biochemistry 30:7014-7020(1991).
RN   [2]
RP   SEQUENCE OF 1-9 FROM N.A.
RC   STRAIN=129/Sv;
RX   MEDLINE=95100931; PubMed=7802637;
RA   Hagihara K., Wu-Peng X.S., Funabashi T., Kato J., Pfaff D.W.;
RT   "Nucleic acid sequence and DNase hypersensitive sites of the 5'
RT   region of the mouse progesterone receptor gene.";
RL   Biochem. Biophys. Res. Commun. 205:1093-1101(1994).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3
CC       subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M68915; AAA39971.1; -.
DR   EMBL; U12644; AAA66067.1; -.
DR   PIR; A39596; A39596.
DR   HSSP; P06401; 1A28.
DR   TRANSPAC; T04680; -.
DR   MGD; MGI:97567; Pgr.
DR   InterPro; IPR000536; Hormone_rec_lig.
DR   InterPro; IPR000128; Progest_receptor.
DR   InterPro; IPR001723; Stdhrmn_receptor.
DR   InterPro; IPR001628; Znf_C4steroid.
DR   Pfam; PF00104; hormone_rec; 1.
DR   Pfam; PF02161; Prog_receptor; 1.
DR   Pfam; PF00105; zf-C4; 1.
DR   PRINTS; PR00398; STRDHORMONER.
DR   PRINTS; PR00047; STROIDFINGER.
DR   ProDom; PD000035; Znf_C4steroid; 1.
DR   SMART; SM00430; HOLI; 1.
DR   SMART; SM00399; ZnF_C4; 1.
DR   PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW   Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW   Zinc-finger; Steroid-binding.
FT   DOMAIN        1    556       MODULATING, PRO-RICH.
FT   DNA_BIND     557    622       NUCLEAR RECEPTOR-TYPE.
FT   ZN_FING      557    577       C4-TYPE.
FT   ZN_FING      593    617       C4-TYPE.
FT   DOMAIN       671    923       STEROID-BINDING.
FT   DOMAIN       184    188       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ   SEQUENCE   923 AA;  99073 MW;  9415F1ED343BEE3F CRC64;

  Query Match             55.3%;  Score 758;  DB 1;  Length 923;
  Best Local Similarity   55.1%;  Pred. No. 6.8e-63;
  Matches  136;  Conservative   56;  Mismatches   55;  Indels    0;  Gaps    0;

Qy         12 PIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFRNLHV 71
              |  :|:| :||| |: ||||| :|| :: ::||:|||:|||||: ||||:|:||||||:
Db        675 PPLINLLMSIEPDVIYAGHDNTKPDTSSSLLTSLNQLGERQLLSVVKWSKSLPGFRNLHI 734

Qy         72 DDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHL 131
              |||: :|||||| |||| :|||| :|: :|||||||| || || :   || |: |  :
Db        735 DDQITLIQYSWMSLMVFGLGWRSYKHVSGQMLYFAPDLILNEQRMKELSFYSLCLTMWQI 794

Qy        132 SQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPT 191
              |||  ||:| :|||||| ||| : ||::||::|  |:|:| :||:|| : |  ::|
Db        795 PQEFVKLQVTHEEFLCMKVLLLLNTIPLEGLRSQSQFEEMRSSYIRELIKAIGLRQKGVV 854

Qy        192 SCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQVPKILSGK 251
              |:|||||||||||:  : ::|| :  :  |:|  ::|:|||||:|:|: |:||||:|
Db        855 PTSQRFYQLTKLLDSLHDLVKQLHLYCLNTFIQSRTLAVEFPEMMSEVIAAQLPKILAGM 914

Qy        252 VKPIYFH 258
              |||: ||
Db        915 VKPLLFH 921


RESULT 14
PRGR_CHICK
ID   PRGR_CHICK     STANDARD;      PRT;   786 AA.
AC   P07812; Q90946;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-AUG-1988 (Rel. 08, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Progesterone receptor (PR).
GN   PGR OR NR3C3.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88166640; PubMed=3443098;
RA   Gronemeyer H., Turcotte B., Quirin-Stricker C., Bocquel M.T.,
RA   Meyer M.E., Krozowski Z., Jeltsch J.M., Lerouge T., Garnier J.M.,
RA   Chambon P.;
RT   "The chicken progesterone receptor: sequence, expression and
RT   functional analysis.";
RL   EMBO J. 6:3985-3994(1987).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91042592; PubMed=3153474;
RA   Conneely O.M., Dobson A.D.W., Tsai M.-J., Beattie W.G., Toft D.O.,
RA   Huckaby C.S., Zarucki T., Schrader W.T., O'Malley B.W.;
RT   "Sequence and expression of a functional chicken progesterone
RT   receptor.";
RL   Mol. Endocrinol. 1:517-525(1987).
RN   [3]
RP   SEQUENCE OF 128-164 FROM N.A.
RX   MEDLINE=86289413; PubMed=2426779;
RA   Conneely O.M., Sullivan W.P., Toft D.O., Birnbaumer M., Cook R.G.,
RA   Maxwell B.L., Zarucki-Schulz T., Greene G.L., Schrader W.T.,
RA   O'Malley B.W.;
RT   "Molecular cloning of the chicken progesterone receptor.";
RL   Science 233:767-770(1986).
RN   [4]
RP   SEQUENCE OF 417-490 FROM N.A.
```

```
RX   MEDLINE=86287271; PubMed=2426697;
RA   Jeltsch J.M., Krozowski Z., Quirin-Stricker C., Gronemeyer H.,
RA   Simpson R.J., Garnier J.M., Krust A., Jacob F., Chambon P.;
RT   "Cloning of the chicken progesterone receptor.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:5424-5428(1986).
RN   [5]
RP   DIFFERENCE BETWEEN FORM 1 AND FORM 2.
RX   MEDLINE=89340509; PubMed=2760059;
RA   Conneely O.M., Kettelberger D.M., Tsai M.-J., Schrader W.T.,
RA   O'Malley B.W.;
RT   "The chicken progesterone receptor A and B isoforms are products of
RT   an alternate translation initiation event.";
RL   J. Biol. Chem. 264:14062-14064(1989).
RN   [6]
RP   SEQUENCE FROM N.A. (ISOFORMS A; A'; B AND B').
RX   MEDLINE=90154085; PubMed=2303488;
RA   Jeltsch J.-M., Turcotte B., Garnier J.-M., Lerouge T., Krozowski Z.,
RA   Gronemeyer H., Chambon P.;
RT   "Characterization of multiple mRNAs originating from the chicken
RT   progesterone receptor gene. Evidence for a specific transcript
RT   encoding form A.";
RL   J. Biol. Chem. 265:3967-3974(1990).
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=4;
CC       Name=A;
CC         IsoId=P07812-1; Sequence=Displayed;
CC       Name=A';
CC         IsoId=P07812-2; Sequence=VSP_003708, VSP_003709;
CC       Name=B;
CC         IsoId=P07812-3; Sequence=VSP_003707;
CC       Name=B';
CC         IsoId=P07812-4; Sequence=VSP_003707, VSP_003708, VSP_003709;
CC   -!- TISSUE SPECIFICITY: OVIDUCT.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3
CC       subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Y00092; CAA68282.1; -.
DR   EMBL; M13972; AAA49034.1; -.
DR   EMBL; M37518; AAA49013.1; -.
DR   EMBL; M37518; AAA49014.1; -.
DR   EMBL; M14278; AAA49035.1; -.
DR   EMBL; M14279; AAA49038.1; -.
DR   EMBL; M14280; AAA49039.1; -.
DR   EMBL; M32732; AAA49011.1; -.
DR   EMBL; M31104; AAA49011.1; JOINED.
DR   EMBL; M32726; AAA49011.1; JOINED.
DR   EMBL; M32727; AAA49011.1; JOINED.
DR   EMBL; M32728; AAA49011.1; JOINED.
DR   EMBL; M32729; AAA49011.1; JOINED.
DR   EMBL; M32730; AAA49011.1; JOINED.
DR   EMBL; M32732; AAA49012.1; -.
DR   EMBL; M31104; AAA49012.1; JOINED.
DR   EMBL; M32726; AAA49012.1; JOINED.
DR   EMBL; M32727; AAA49012.1; JOINED.
DR   EMBL; M32728; AAA49012.1; JOINED.
DR   EMBL; M32729; AAA49012.1; JOINED.
DR   EMBL; M32730; AAA49012.1; JOINED.
DR   EMBL; M31104; AAA49009.1; -.
DR   EMBL; M31104; AAA49010.1; -.
DR   PIR; A35466; A35466.
DR   HSSP; P06401; 1A28.
DR   TRANSFAC; T00698; -.
DR   InterPro; IPR000536; Hormone_rec_lig.
DR   InterPro; IPR000128; Progest_receptor.
DR   InterPro; IPR001723; Stdhrmn_receptor.
DR   InterPro; IPR001628; Znf_C4steroid.
DR   Pfam; PF00104; hormone_rec; 1.
DR   Pfam; PF02161; Prog_receptor; 1.
DR   Pfam; PF00105; zf-C4; 1.
DR   PRINTS; PR00398; STRDHORMONER.
DR   PRINTS; PR00047; STROIDFINGER.
DR   ProDom; PD000035; Znf_C4steroid; 1.
DR   SMART; SM00430; HOLI; 1.
DR   SMART; SM00399; ZnF_C4; 1.
DR   PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW   Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW   Zinc-finger; Steroid-binding; Alternative splicing.
FT   DOMAIN        1    420       MODULATING, PRO-RICH.
FT   DNA_BIND    421    486       NUCLEAR RECEPTOR-TYPE.
FT   ZN_FING     421    441       C4-TYPE.
FT   ZN_FING     457    481       C4-TYPE.
FT   DOMAIN      487    786       STEROID-BINDING.
FT   DOMAIN       48     80       ASP/GLU-RICH (ACIDIC).
FT   VARSPLIC      1    127       Missing (in isoform B and isoform B').
FT                                /FTId=VSP_003707.
FT   VARSPLIC    452    458       QHNYLCA -> TISYHCS (in isoform A' and
FT                                isoform B').
FT                                /FTId=VSP_003708.
FT   VARSPLIC    459    786       Missing (in isoform A' and isoform B').
FT                                /FTId=VSP_003709.
FT   CONFLICT     58     58       E -> DD (IN REF. 2).
FT   CONFLICT    480    480       K -> N (IN REF. 2).
FT   CONFLICT    489    489       G -> A (IN REF. 2).
FT   CONFLICT    577    577       R -> T (IN REF. 2).
FT   CONFLICT    642    642       M -> I (IN REF. 2).
SQ   SEQUENCE   786 AA;  85743 MW;  659559950BC45ED9 CRC64;

  Query Match           54.9%;  Score 752;  DB 1;  Length 786;
  Best Local Similarity 56.3%;  Pred. No. 2e-62;
  Matches  139;  Conservative   53;  Mismatches   55;  Indels    0;  Gaps    0;

Qy        12 PIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFRNLHV 71
             |  ::||  ||| || ||:|| :|:: ::||:||| | ||||: |||||:| ||||||||:
Db       538 PPMISVLRGIEPEVVYAGYDNTKPETPSSLLTSLNHLCERQLLCVVKWSKLLPGFRNLHI 597

Qy        72 DDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHL 131
             |||: :|||||| ||||||||||: :|: :|||||||: || || :|  || |: |  |
Db       598 DDQITLIQYSWMSLMVFAMGWRSYKHVSGQMLYFAPDLILNEQRMKESSFYSLCLSMWQL 657

Qy       132 SQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPT 191
             |||  ||:: :|||||||||| : ||::||::|  |||:| :||:|| : |  ::|
Db       658 PQEFVRLQVSQEEFLCMKALLLLNTIPLEGLRSQSQFDEMRTSYIRELVKAIGLRQKGVV 717

Qy       192 SCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQVPKILSGK 251
             : |:|||||||:||:  : ::|| |  :  ::|  :||:|||||:|:|: |:||||:|
Db       718 ANSQRFYQLTKLMDSMHDLVKQLHLFCLNTFLQSRALSVEFPEMMSEVIAAQLPKILAGM 777

Qy       252 VKPIYFH 258
             |||: ||
Db       778 VKPLLFH 784


RESULT 15
PRGR_SHEEP
ID   PRGR_SHEEP     STANDARD;      PRT;   377 AA.
AC   Q28590;
DT   01-NOV-1997 (Rel. 35, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Progesterone receptor (PR) (Fragment).
GN   PGR OR NR3C3.
OS   Ovis aries (Sheep).
```

```
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Romanov; TISSUE=Uterus;
RA   Madigou T., Tiffoche C., le Gal F., Pelletier J., Thieulant M.;
RL   Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC       THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC       PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC       A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC   -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3
CC       subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z66555; CAA91447.1; -.
DR   HSSP; P06401; 1A28.
DR   InterPro; IPR000536; Hormone_rec_lig.
DR   InterPro; IPR000128; Progest_receptor.
DR   InterPro; IPR001723; Stdhrmn_receptor.
DR   InterPro; IPR001628; Znf_C4steroid.
DR   Pfam; PF00104; hormone_rec; 1.
DR   Pfam; PF02161; Prog_receptor; 1.
DR   Pfam; PF00105; zf-C4; 1.
DR   PRINTS; PR00398; STRDHORMONER.
DR   PRINTS; PR00047; STROIDFINGER.
DR   ProDom; PD000035; Znf_C4steroid; 1.
DR   SMART; SM00430; HOLI; 1.
DR   SMART; SM00399; ZnF_C4; 1.
DR   PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW   Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW   Zinc-finger; Steroid-binding.
FT   NON_TER       1      1
FT   DOMAIN       <1     15       MODULATING, PRO-RICH.
FT   DNA_BIND      18     83       NUCLEAR RECEPTOR-TYPE.
FT   ZN_FING       18     38       C4-TYPE.
FT   ZN_FING       54     78       C4-TYPE.
FT   DOMAIN       128   >377       STEROID-BINDING.
FT   NON_TER      377    377
SQ   SEQUENCE   377 AA;  42904 MW;  3141B65587F7493C CRC64;

  Query Match             53.8%;  Score 738;  DB 1;  Length 377;
  Best Local Similarity   55.0%;  Pred. No. 1.6e-61;
  Matches  133;  Conservative   57;  Mismatches   52;  Indels    0;  Gaps    0;

Qy         12 PIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFRNLHV 71
              |  :|:| :||| :| ||||::||: ::||:|||:|||||: ||||:|:||||||:
Db        136 PPLINLLMSIEPDMVYAGHDNSKPDTSSSLLTSLNQLGERQLLSVVKWSKSLPGFRNLHI 195

Qy         72 DDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHL 131
              |||: :|||||| |||| :||||: :|: :||||||||: || || :| || |: |  :
Db        196 DDQITLIQYSWMSLMVFGLGWRSYKHVSGQMLYFAPDLILNEQRMKESSFYSLCLTMWQI 255

Qy        132 SQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPT 191
              |||  ||:: :|||||| ||| : ||::||::|  |:|:| :||:|| : |  ::|
Db        256 PQEFVKLQVSQEEFLCMKVLLLLNTIPLEGLRSQNQFEEMRSSYIRELIKAIGLRQKGVV 315

Qy        192 SCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQVPKILSGK 251
              |:||||||||||::  : ::|| :  :  |:|  :||:|||||:|:|: |:||||:|
Db        316 PSSQRFYQLTKLLDNLHDLVKQLHLYCLNTFIQSRALSVEFPEMMSEVIAAQLPKILAGM 375

Qy        252 VK 253
              ||
Db        376 VK 377
```

Search completed: October 9, 2003, 11:41:31
Job time : 12 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 9, 2003, 11:40:28 ; Search time 19 Seconds
(without alignments)
1315.992 Million cell updates/sec

Title: US-09-687-609A-1
Perfect score: 1371
Sequence: 1 GSHMIEGYECQPIFLNVLEA.....SVQVPKILSGKVKPIYFHTQ 260

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1349.5	98.4	899	2	A35895	androgen receptor
2	1349.5	98.4	902	2	B40494	androgen receptor
3	1349.5	98.4	919	2	A39248	androgen receptor
4	1341.5	97.8	910	2	A34721	androgen receptor
5	1306	95.3	911	2	B34721	androgen receptor
6	1217	88.8	344	2	I51330	androgen receptor
7	900	65.6	848	2	JG0194	androgen receptor
8	766	55.9	930	2	A25923	progesterone recep
9	765	55.8	933	1	QRHUP	progesterone recep
10	759	55.4	923	2	I53280	progesterone recep
11	758	55.3	923	2	A39596	progesterone recep
12	752	54.9	786	2	A35466	progesterone recep
13	665	48.5	984	2	A29513	mineralocorticoid
14	662.5	48.3	777	1	QRHUGA	glucocorticoid rec
15	662	48.3	981	2	A41401	mineralocorticoid
16	657.5	48.0	783	1	A25691	glucocorticoid rec
17	650.5	47.4	776	1	S44047	glucocorticoid rec
18	648.5	47.3	758	2	S60586	glucocorticoid rec
19	644.5	47.0	771	2	A54273	glucocorticoid rec
20	638.5	46.6	795	1	QRRTG	glucocorticoid rec
21	564	41.1	742	1	QRHUGB	glucocorticoid rec
22	387	28.2	166	2	S35795	androgen receptor
23	286	20.9	110	2	I53287	progesterone recep
24	222	16.2	535	2	S58224	oestrogen receptor
25	218.5	15.9	589	1	QRCHE	estrogen receptor
26	218	15.9	586	1	QRXLE	estrogen receptor
27	216.5	15.8	595	2	I47140	estradiol receptor
28	215	15.7	620	2	T10423	estrogen receptor
29	214.5	15.6	574	2	A37197	estrogen receptor
30	213.5	15.6	599	1	QRMSE	estrogen receptor
31	211.5	15.4	600	1	QRRTE	estrogen receptor
32	207.5	15.1	503	2	JW0046	estrogen receptor
33	204.5	14.9	595	1	QRHUE	estrogen receptor
34	204.5	14.9	701	2	S64737	80K estrogen recep
35	203.5	14.8	477	2	S71400	estrogen receptor
36	203.5	14.8	530	2	JC5939	estrogen receptor
37	197	14.4	462	2	S44490	RXR type hormone r
38	171	12.5	521	2	A29345	steroid hormone re
39	170	12.4	433	2	B29345	steroid hormone re
40	168	12.3	433	2	S58087	estrogen receptor
41	148.5	10.8	461	1	A46077	steroid hormone re
42	148	10.8	808	1	S33708	nuclear steroid/th
43	145.5	10.6	462	2	A40716	steroid hormone re
44	145	10.6	462	1	A56120	steroid hormone re
45	143	10.4	463	2	S26670	retinoic acid rece

ALIGNMENTS

RESULT 1
A35895
androgen receptor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 31-Oct-1990 #sequence_revision 31-Oct-1990 #text_change 20-Sep-1999
C;Accession: A35895; A37255; A37908; S34398; S17198; S40626; I49501; S12082
R;He, W.W.; Fischer, L.M.; Sun, S.; Bilhartz, D.L.; Zhu, X.; Young, C.Y.F.; Kelley, D.B.;
Biochem. Biophys. Res. Commun. 171, 697-704, 1990
A;Title: Molecular cloning of androgen receptors from divergent species with a polymerase
eptor cDNA probes from dog, guinea pig and clawed frog.
A;Reference number: A35895; MUID:90386642; PMID:2403358
A;Accession: A35895
A;Molecule type: mRNA
A;Residues: 1-899 <HEA>
A;Cross-references: EMBL:X53779; NID:g49966; PIDN:CAA37795.1; PID:g49967
R;Gaspar, M.L.; Meo, T.; Tosi, M.
Mol. Endocrinol. 4, 1600-1610, 1990
A;Title: Structure and size distribution of the androgen receptor mRNA in wild-type and
A;Reference number: A37255; MUID:91133433; PMID:2178222
A;Accession: A37255
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-899 <GAS>
A;Cross-references: GB:M37890; NID:g191935; PIDN:AAA37234.1; PID:g191936
R;Charest, N.J.; Zhou, Z.; Lubahn, D.B.; Olsen, K.L.; Wilson, E.M.; French, F.S.
Mol. Endocrinol. 5, 573-581, 1991
A;Title: A frameshift mutation destabilizes androgen receptor messenger RNA in the Tfm mo
A;Reference number: A37908; MUID:92017874; PMID:1681426
A;Accession: A37908
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-899 <CHA>
A;Cross-references: GB:S56585; NID:g236048; PIDN:AAB19916.1; PID:g236049
R;Faber, P.W.; King, A.; van Rooij, H.C.J.; Brinkmann, A.O.; de Both, N.J.; Trapman, J.
Biochem. J. 278, 269-278, 1991
A;Title: The mouse androgen receptor: Functional analysis of the protein and characteriza
A;Reference number: S17198; MUID:91354214; PMID:1883336
A;Accession: S34398
A;Molecule type: DNA
A;Residues: 1-899 <FAB>
A;Cross-references: EMBL:X59592; NID:g49968; PIDN:CAA42160.1; PID:g49969
A;Accession: S17198
A;Molecule type: mRNA
A;Residues: 1-899 <FA2>
A;Cross-references: GB:X59592; NID:g49968; PIDN:CAA42160.1; PID:g49969
R;He, W.W.; Kumar, M.V.; Tindall, D.J.
Nucleic Acids Res. 19, 2373-2378, 1991
A;Title: A frame-shift mutation in the androgen receptor gene causes complete androgen in
A;Reference number: S40626; MUID:91252278; PMID:2041777
A;Accession: S40626
A;Status: preliminary
A;Molecule type: mRNA

A;Residues: 366-413 <HEW>
A;Cross-references: EMBL:X53779
R;Gaspar, M.
Proc. Natl. Acad. Sci. U.S.A. 88, 8606-8610, 1991
A;Title: A single base deletion in the Tfm androgen receptor gene creates a short-lived
A;Reference number: I49501; MUID:92020902; PMID:1924321
A;Accession: I49501
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-899 <RES>
A;Cross-references: GB:M37890; NID:g191935; PIDN:AAA37234.1; PID:g191936
C;Genetics:
A;Introns: 518/2; 569/1; 608/1; 704/1; 752/2; 796/1; 848/3
C;Superfamily: unassigned erbA-related proteins; erbA transforming protein homology
C;Keywords: DNA binding; transcription regulation; zinc finger
F;1-536/Domain: regulatory #status predicted <REG>
F;537-795/Domain: erbA transforming protein homology <ERBA>
F;539-559/Region: zinc finger
F;575-599/Region: zinc finger
F;650-899/Domain: hormone binding #status predicted <LIG>

Query Match 98.4%; Score 1349.5; DB 2; Length 899;
Best Local Similarity 99.6%; Pred. No. 2.5e-118;
Matches 258; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        642 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 700

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        701 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 760

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        761 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 820

Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        821 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 880

Qy        242 VQVPKILSGKVKPIYFHTQ 260
              ||||||||||||||||||
Db        881 VQVPKILSGKVKPIYFHTQ 899
```

RESULT 2
B40494
androgen receptor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 20-Sep-1999
C;Accession: B40494; A34943; A36283; B40108
R;Chang, C.; Kokontis, J.; Liao, S.
Proc. Natl. Acad. Sci. U.S.A. 85, 7211-7215, 1988
A;Title: Structural analysis of complementary DNA and amino acid sequences of human and
A;Reference number: A40494; MUID:89017168; PMID:3174628
A;Accession: B40494
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-902 <CHA>
A;Cross-references: GB:M23264; NID:g202967; PIDN:AAA40759.1; PID:g202968
R;Tan, J.; Joseph, D.R.; Quarmby, V.E.; Lubahn, D.B.; Sar, M.; French, F.S.; Wilson, E.M
Mol. Endocrinol. 2, 1276-1285, 1988
A;Title: The rat androgen receptor: primary structure, autoregulation of its messenger r
A;Reference number: A34943; MUID:89112209; PMID:3216867
A;Accession: A34943
A;Molecule type: mRNA
A;Residues: 1-388,'S',390-902 <TAN>
A;Cross-references: GB:M20133; NID:g202895; PIDN:AAA40733.1; PID:g202896
R;Yarbrough, W.G.; Quarmby, V.E.; Simental, J.A.; Joseph, D.R.; Sar, M.; Lubahn, D.B.; C
J. Biol. Chem. 265, 8893-8900, 1990
A;Title: A single base mutation in the androgen receptor gene causes androgen insensitiv
A;Reference number: A36283; MUID:90256822; PMID:2341409
A;Accession: A36283
A;Molecule type: mRNA
A;Residues: 1-194,196-902 <YAR>
A;Cross-references: GB:J05454
R;Chang, C.; Kokontis, J.; Liao, S.
Science 240, 324-326, 1988
A;Title: Molecular cloning of human and rat complementary DNA encoding androgen receptor
A;Reference number: A40108; MUID:88178111; PMID:3353726
A;Accession: B40108
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 540-611 <CH2>
C;Superfamily: unassigned erbA-related proteins; erbA transforming protein homology
C;Keywords: DNA binding; zinc finger
F;540-798/Domain: erbA transforming protein homology <ERBA>
F;542-562/Region: zinc finger
F;578-602/Region: zinc finger

Query Match 98.4%; Score 1349.5; DB 2; Length 902;
Best Local Similarity 99.6%; Pred. No. 2.5e-118;
Matches 258; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        645 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 703

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        704 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 763

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        764 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 823

Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        824 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 883

Qy        242 VQVPKILSGKVKPIYFHTQ 260
              ||||||||||||||||||
Db        884 VQVPKILSGKVKPIYFHTQ 902
```

RESULT 3
A39248
androgen receptor - human
C;Species: Homo sapiens (man)
C;Date: 04-Oct-1991 #sequence_revision 04-Oct-1991 #text_change 24-Nov-1999
C;Accession: A39248; A30328; A40109; A60946; A34942; A27653; A40108; A40494; A32224; A40
R;Lubahn, D.B.; Brown, T.R.; Simental, J.A.; Higgs, H.N.; Migeon, C.J.; Wilson, E.M.; Fr
Proc. Natl. Acad. Sci. U.S.A. 86, 9534-9538, 1989
A;Title: Sequence of the intron/exon junctions of the coding region of the human androge
A;Reference number: A39248; MUID:90083302; PMID:2594783
A;Accession: A39248
A;Molecule type: DNA
A;Residues: 1-919 <LUB>
A;Cross-references: GB:M27423; GB:M27430; NID:g178904; PIDN:AAA51886.1; PID:g178906
R;Faber, P.W.; Kuiper, G.G.J.M.; van Rooij, H.C.J.; van der Korput, J.A.G.M.; Brinkmann,
Mol. Cell. Endocrinol. 61, 257-262, 1989
A;Title: The N-terminal domain of the human androgen receptor is encoded by one, large e
A;Reference number: A30328; MUID:89137730; PMID:2917688
A;Accession: A30328
A;Molecule type: DNA
A;Residues: 1-77,79-165,'A',167-389,'L',391-464,473-538 <FAB>
A;Cross-references: GB:M20260
R;Lubahn, D.B.; Joseph, D.R.; Sullivan, P.M.; Willard, H.F.; French, F.S.; Wilson, E.M.
Science 240, 327-330, 1988
A;Title: Cloning of human androgen receptor complementary DNA and localization to the X
A;Reference number: A40109; MUID:88178112; PMID:3353727
A;Accession: A40109
A;Molecule type: DNA
A;Residues: 559-624 <LU2>
A;Cross-references: GB:M20132

R;Kuiper, G.G.J.M.; Faber, P.W.; van Rooij, H.C.J.; van der Korput, J.A.G.M.; Ris-Stalpe
J. Mol. Endocrinol. 2, R1-R4, 1989
A;Title: Structural organization of the human androgen receptor gene.
A;Reference number: A60946; MUID:89322749; PMID:2546571
A;Accession: A60946
A;Molecule type: DNA
A;Residues: 536-540;587-591;626-631;723-726;770-774;814-818;867-870 <KUI>
R;Lubahn, D.B.; Joseph, D.R.; Sar, M.; Tan, J.; Higgs, H.N.; Larson, R.E.; French, F.S.;
Mol. Endocrinol. 2, 1265-1275, 1988
A;Title: The human androgen receptor: complementary deoxyribonucleic acid cloning, seque
A;Reference number: A34942; MUID:89112208; PMID:3216866
A;Accession: A34942
A;Molecule type: mRNA
A;Residues: 1-919 <LU3>
A;Cross-references: GB:M20132; NID:g178627; PIDN:AAA51729.1; PID:g178628; GB:J03180
R;Trapman, J.; Klaassen, P.; Kuiper, G.G.J.M.; van der Korput, J.A.G.M.; Faber, P.W.; va
Biochem. Biophys. Res. Commun. 153, 241-248, 1988
A;Title: Cloning, structure and expression of a cDNA encoding the human androgen receptc
A;Reference number: A27653; MUID:88240407; PMID:3377788
A;Accession: A27653
A;Molecule type: mRNA
A;Residues: 468-564,'K',566-919 <TRA>
A;Cross-references: GB:M20260; NID:g178891; PIDN:AAA51774.1; PID:g178892
A;Note: the authors translated the codon AAG for residue 565 as Glu
R;Chang, C.; Kokontis, J.; Liao, S.
Science 240, 324-326, 1988
A;Title: Molecular cloning of human and rat complementary DNA encoding androgen receptor
A;Reference number: A40108; MUID:88178111; PMID:3353726
A;Accession: A40108
A;Molecule type: mRNA
A;Residues: 557-628 <CHA>
A;Cross-references: GB:M18624
R;Chang, C.; Kokontis, J.; Liao, S.
Proc. Natl. Acad. Sci. U.S.A. 85, 7211-7215, 1988
A;Title: Structural analysis of complementary DNA and amino acid sequences of human and
A;Reference number: A40494; MUID:89017168; PMID:3174628
A;Accession: A40494
A;Molecule type: mRNA
A;Residues: 1-74,79-89,'H',90-472,'GGG',473-474,'E',476-644,'N',646-919 <CH2>
A;Cross-references: GB:M23263
R;Tilley, W.D.; Marcelli, M.; Wilson, J.D.; McPhaul, M.J.
Proc. Natl. Acad. Sci. U.S.A. 86, 327-331, 1989
A;Title: Characterization and expression of a cDNA encoding the human androgen receptor.
A;Reference number: A32224; MUID:89098909; PMID:2911578
A;Accession: A32224
A;Molecule type: mRNA
A;Residues: 1-77,79-211,'R',213-471,473-919 <TIL>
A;Cross-references: GB:M21748; GB:J04150; NID:g178871; PIDN:AAA51771.1; PID:g178872
R;Mowszowicz, I.; Lee, H.J.; Chen, H.T.; Mestayer, C.; Portois, M.C.; Cabrol, S.; Mauvai
Mol. Endocrinol. 7, 861-869, 1993
A;Title: A point mutation in the second zinc finger of the DNA-binding domain of the and
stance.
A;Reference number: A40715; MUID:94019395; PMID:8413310
A;Accession: A40715
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 557-614,'H',616-624 <MOW>
A;Cross-references: PIDN:AAB28340.1; PID:g425580
C;Genetics:
A;Gene: GDB:AR
A;Cross-references: GDB:120556; OMIM:313700
A;Map position: Xq11-Xq12
A;Introns: 538/2; 589/1; 628/1; 724/1; 772/2; 816/1; 868/3
C;Superfamily: unassigned erbA-related proteins; erbA transforming protein homology
C;Keywords: DNA binding; steroid binding; transcription regulation; zinc finger
F;557-815/Domain: erbA transforming protein homology <ERBA>
F;559-579/Region: zinc finger
F;595-619/Region: zinc finger

Query Match 98.4%; Score 1349.5; DB 2; Length 919;
Best Local Similarity 99.6%; Pred. No. 2.6e-118;
Matches 258; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 720

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 780

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 840

Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 900

Qy        242 VQVPKILSGKVKPIYFHTQ 260
              |||||||||||||||||||
Db        901 VQVPKILSGKVKPIYFHTQ 919
```

RESULT 4
A34721
androgen receptor A - human
C;Species: Homo sapiens (man)
C;Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 12-Sep-1997
C;Accession: A34721
R;Govindan, M.V.
Mol. Endocrinol. 4, 417-427, 1990
A;Title: Specific region in hormone binding domain is essential for hormone binding and
A;Reference number: A34721; MUID:90258935; PMID:2342476
A;Accession: A34721
A;Molecule type: mRNA
A;Residues: 1-910 <GOV>
C;Superfamily: unassigned erbA-related proteins; erbA transforming protein homology
C;Keywords: zinc finger
F;548-806/Domain: erbA transforming protein homology <ERBA>
F;550-570/Region: zinc finger
F;586-610/Region: zinc finger

Query Match 97.8%; Score 1341.5; DB 2; Length 910;
Best Local Similarity 98.8%; Pred. No. 1.4e-117;
Matches 256; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

```
Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 711

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        712 ALPGLRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 771

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        772 YSQCVRMRHLSQEFGWLQITPQEFLCMKAMLLFSIIPVDGLKNQKFFDELRMNYIKELDR 831

Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        832 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 891

Qy        242 VQVPKILSGKVKPIYFHTQ 260
              |||||||||||||||||||
Db        892 VQVPKILSGKVKPIYFHTQ 910
```

RESULT 5
B34721
androgen receptor B - human
C;Species: Homo sapiens (man)
C;Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 12-Sep-1997
C;Accession: B34721
R;Govindan, M.V.
Mol. Endocrinol. 4, 417-427, 1990

```
A;Title: Specific region in hormone binding domain is essential for hormone binding and
A;Reference number: A34721; MUID:90258935; PMID:2342476
A;Accession: B34721
A;Molecule type: mRNA
A;Residues: 1-911 <GO2>
C;Superfamily: unassigned erbA-related proteins; erbA transforming protein homology
C;Keywords: zinc finger
F;548-806/Domain: erbA transforming protein homology <ERBA>
F;550-570/Region: zinc finger
F;586-610/Region: zinc finger

  Query Match             95.3%;  Score 1306;  DB 2;  Length 911;
  Best Local Similarity   97.3%;  Pred. No. 3.1e-114;
  Matches  253;  Conservative    1;  Mismatches    4;  Indels    2;  Gaps    2;

Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 711

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        712 ALPGLRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 771

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQ-KFFDELRMNYIKELD 180
              ||||||||||||||||||||||||||||:|||||||||||||||  ||  ||||||||||
Db        772 YSQCVRMRHLSQEFGWLQITPQEFLCMKAMLLFSIIPVDGLKNQIFFFFLLRMNYIKELD 831

Qy        181 RIIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEII 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        832 RIIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEII 891

Qy        241 SVQVPKILSGKVKPIYFHTQ 260
              ||||||||||||||||||||
Db        892 SVQVPKILSGKVKPIYFHTQ 911


RESULT 6
I51330
androgen receptor - common canary (fragment)
C;Species: Serinus canaria (common canary)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 20-Sep-1999
C;Accession: I51330
R;Nastiuk, K.L.; Clayton, D.F.
Endocrinology 134, 640-649, 1994
A;Title: Seasonal and tissue-specific regulation of canary androgen receptor messenger r
A;Reference number: I51330; MUID:94130808; PMID:8299561
A;Accession: I51330
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-344 <NAS>
A;Cross-references: GB:L25901; NID:g414733; PIDN:AAA17402.1; PID:g414734
C;Genetics:
A;Gene: AR
C;Superfamily: unassigned erbA-related proteins; erbA transforming protein homology
C;Keywords: zinc finger
F;1-251/Domain: erbA transforming protein homology (fragment) <ERBA>

  Query Match             88.8%;  Score 1217;  DB 2;  Length 344;
  Best Local Similarity   92.7%;  Pred. No. 2e-106;
  Matches  227;  Conservative   13;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          5 IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALP 64
              |:|||||||||||||||||||||||:|||||: ||:|||||||||||:|||||||||
Db        100 IDGYECQPIFLNVLEAIEPGVVCAGHDNSQPDSPSNLLTSLNELGERQLVYVVKWAKALP 159

Qy         65 GFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQ 124
              ||||||||||::|||||||||||||||||||||||||||||||||||||||||||||||
Db        160 GFRNLHVDDQMSIIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQ 219

Qy        125 CVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIA 184
              |:|||||||||||||| |||||||| |||||||||||||| ||||||||||||||||||
Db        220 CIRMRHLSQEFGWLQITPQGFLCMKALLFFSIIPVDGLKNQKLFDELRMNYIKELDRIIA 279

Qy        185 CKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQV 244
              ||||||||||||||||||:|||| |||::|||||||||||:||||||:||||||||||||
Db        280 CKRKNPTSCSRRFYQLTKVLDSVTPIAKDLHQFTFDLLIKAHMVSVDYPEMMAEIISVQV 339

Qy        245 PKILS 249
              |||||
Db        340 PKILS 344


RESULT 7
JG0194
androgen receptor - Japanese eel
C;Species: Anguilla japonica (Japanese eel)
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 11-May-2000
C;Accession: JG0194
R;Todo, T.; Ikeuchi, T.; Kobayashi, T.; Nagahama, Y.
Biochem. Biophys. Res. Commun. 254, 378-383, 1999
A;Title: Fish androgen receptor: cDNA cloning, steroid activation of transcription in tr
A;Reference number: JG0194; MUID:99119319; PMID:9918846
A;Accession: JG0194
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-848 <TOD>
C;Superfamily: unassigned erbA-related proteins; erbA transforming protein homology
F;491-745/Domain: erbA transforming protein homology <ERB>

  Query Match             65.6%;  Score 900;  DB 2;  Length 848;
  Best Local Similarity   68.1%;  Pred. No. 3.7e-76;
  Matches  173;  Conservative   35;  Mismatches   42;  Indels    4;  Gaps    2;

Qy          7 GYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGF 66
              |:  | :|||:|||||| || ||||  |||| |:||:|||||||||| |||||| :|||
Db        596 GFHTQSMFLNILEAIEPEVVNAGHDYGQPDSAASLLTSLNELGERQLVKVVKWAKGMPGF 655

Qy         67 RNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCV 126
              |:|:||||| |||   || :||||:||||| || ||||||||||||:||  | ||  |:
Db        656 RSLYVDDQMTVIQKHWMAVMVFALGWRSFKNVKSRMLYFAPDLVFNEHRMQVSTMYEHCI 715

Qy        127 RMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACK 186
              ||:: ||||  ||:: :|||||||||||| |||:||| | |||||| :|| ||||::: :
Db        716 RMKNFSQEFAMLQVSQEEFLCMKALLLFSTIPVEGLKGQNFFDELRRSYINELDRLVSFR 775

Qy        187 RKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVS--VDFPEMMAEIISVQV 244
              |  :||| || |||:||||:||: ::|||||||| ::|  :|  | ||||::||||| |
Db        776 SK--SSCSERFQQLTRLLDSLQPVLKKLHQFTFDLFVQSQNLSNQVCFPEMISEIISVHV 833

Qy        245 PKILSGKVKPIYFH 258
              ||||:| |||| ||
Db        834 PKILAGTVKPILFH 847


RESULT 8
A25923
progesterone receptor - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 16-Aug-1988 #sequence_revision 16-Aug-1988 #text_change 20-Aug-1999
C;Accession: A25923
R;Loosfelt, H.; Atger, M.; Misrahi, M.; Guiochon-Mantel, A.; Meriel, C.; Logeat, F.; Ben
Proc. Natl. Acad. Sci. U.S.A. 83, 9045-9049, 1986
A;Title: Cloning and sequence analysis of rabbit progesterone-receptor complementary DNA
A;Reference number: A25923; MUID:87067449; PMID:3538016
A;Accession: A25923
A;Molecule type: mRNA
A;Residues: 1-930 <LOO>
A;Cross-references: GB:M14547; NID:g165631; PIDN:AAA31443.1; PID:g165632
C;Superfamily: progesterone receptor; erbA transforming protein homology
C;Keywords: DNA binding; nucleus; steroid hormone receptor; transcription regulation; zi
F;566-826/Domain: erbA transforming protein homology <ERBA>
F;568-588/Region: zinc finger
F;604-628/Region: zinc finger

  Query Match             55.9%;  Score 766;  DB 2;  Length 930;
  Best Local Similarity   55.5%;  Pred. No. 1.6e-63;
```

```
  Matches  137;  Conservative   57;  Mismatches   53;  Indels    0;  Gaps    0;

Qy          12 PIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFRNLHV 71
               |  :|:| :||| |: ||||| :||: ::||:|||:|||||: ||||:|:||||||||:
Db         682 PPLINLLMSIEPDVIYAGHDNTKPDTSSSLLTSLNQLGERQLLSVVKWSKSLPGFRNLHI 741

Qy          72 DDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHL 131
               |||: :|||||| |||| :||||: :|: :||||||||: || || :|  || |: |  :
Db         742 DDQITLIQYSWMSLMVFGLGWRSYKHVSGQMLYFAPDLILNEQRMKESSFYSLCLTMWQI 801

Qy         132 SQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPT 191
               |||  ||:: :|||||| ||| : ||::||:|  |:|:| :||:||| : |  ::|
Db         802 PQEFVKLQVSQEEFLCMKVLLLLNTIPLEGLRSQSQFEEMRSSYIRELIKAIGLRQKGVV 861

Qy         192 SCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQVPKILSGK 251
               | |:||||||||||::  : ::|| :  :  |:|  :||:|||||:|:|: |:||||:|
Db         862 SSSQRFYQLTKLLDNLHDLVKQLHLYCLNTFIQSRALSVEFPEMMSEVIAAQLPKILAGM 921

Qy         252 VKPIYFH 258
               |||: ||
Db         922 VKPLLFH 928


RESULT 9
QRHUP
progesterone receptor form B - human
N;Alternate names: hPR
N;Contains: progesterone receptor form A
C;Species: Homo sapiens (man)
C;Date: 30-Jun-1987 #sequence_revision 18-Nov-1994 #text_change 22-Jun-1999
C;Accession: S09971; S12464; A03245
R;Kastner, P.; Krust, A.; Turcotte, B.; Stropp, U.; Tora, L.; Gronemeyer, H.; Chambon, P
EMBO J. 9, 1603-1614, 1990
A;Title: Two distinct estrogen-regulated promoters generate transcripts encoding the two
A;Reference number: S09971; MUID:90228361; PMID:2328727
A;Accession: S09971
A;Molecule type: mRNA
A;Residues: 1-933 <KAS>
A;Cross-references: EMBL:X51730
R;Kastner, P.
submitted to the EMBL Data Library, February 1990
A;Reference number: S12464
A;Accession: S12464
A;Molecule type: mRNA
A;Residues: 1-343,'T',345-933 <KA2>
A;Cross-references: EMBL:X51730; NID:g35651; PIDN:CAA36018.1; PID:g35652
R;Misrahi, M.; Atger, M.; D'Auriol, L.; Loosfelt, H.; Meriel, C.; Fridlansky, F.; Guioch
Biochem. Biophys. Res. Commun. 143, 740-748, 1987
A;Title: Complete amino acid sequence of the human progesterone receptor deduced from cl
A;Reference number: A03245; MUID:87184565; PMID:3551956
A;Accession: A03245
A;Molecule type: mRNA
A;Residues: 1-225,'G',227-255,'V',257-659,'V',661-933 <MIS>
A;Cross-references: GB:M15716; NID:g189934; PIDN:AAA60081.1; PID:g189935
C;Genetics:
A;Gene: GDB:PGR
A;Cross-references: GDB:119493; OMIM:264080
A;Map position: 11q22.1-11q22.3
C;Superfamily: progesterone receptor; erbA transforming protein homology
C;Keywords: alternative splicing; DNA binding; nucleus; phosphoprotein; steroid hormone
F;1-933/Product: progesterone receptor form B #status predicted <MA1>
F;165-933/Product: progesterone receptor form A #status predicted <MA2>
F;565-829/Domain: erbA transforming protein homology <ERBA>
F;567-587/Region: zinc finger CCCC motif
F;603-627/Region: zinc finger CCCC motif
F;681-933/Domain: steroid binding #status predicted <STB>
F;41/Binding site: phosphate (Thr) (covalent) #status predicted
F;227,232,552,793/Binding site: phosphate (Ser) (covalent) #status predicted
F;329,374,601/Binding site: phosphate (Tyr) (covalent) #status predicted

  Query Match           55.8%;  Score 765;  DB 1;  Length 933;
  Best Local Similarity  55.5%;  Pred. No. 2e-63;
  Matches  137;  Conservative   57;  Mismatches   53;  Indels    0;  Gaps    0;

Qy          12 PIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFRNLHV 71
               |  :|:| :||| || ||||| :||: ::||:|||:|||||: ||||:|:||||||||:
Db         685 PPLINLLMSIEPDVIYAGHDNTKPDTSSSLLTSLNQLGERQLLSVVKWSKSLPGFRNLHI 744

Qy          72 DDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHL 131
               |||: :|||||| |||| :||||: :|: :||||||||: || || :|  || |: |  :
Db         745 DDQITLIQYSWMSLMVFGLGWRSYKHVSGQMLYFAPDLILNEQRMKESSFYSLCLTMWQI 804

Qy         132 SQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPT 191
               |||  ||:: :|||||| ||| : ||::||:|  |:|:| :||:||| : |  ::|
Db         805 PQEFVKLQVSQEEFLCMKVLLLLNTIPLEGLRSQTQFEEMRSSYIRELIKAIGLRQKGVV 864

Qy         192 SCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQVPKILSGK 251
               | |:||||||||||::  : ::|| :  :  |:|  :||:|||||:|:|: |:||||:|
Db         865 SSSQRFYQLTKLLDNLHDLVKQLHLYCLNTFIQSRALSVEFPEMMSEVIAAQLPKILAGM 924

Qy         252 VKPIYFH 258
               |||: ||
Db         925 VKPLLFH 931


RESULT 10
I53280
progesterone receptor B form - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 20-Aug-1999
C;Accession: I53280; B23733; A49574; A23733
R;Park-Sarge, O.K.; Mayo, K.E.
Endocrinology 134, 709-718, 1994
A;Title: Regulation of the progesterone receptor gene by gonadotropins and cyclic adenos
A;Reference number: I53280; MUID:94130817; PMID:8299566
A;Accession: I53280
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-923 <PAR1>
A;Cross-references: GB:L16922; NID:g463282; PIDN:AAA19916.1; PID:g463283
R;Park, O.K.; Mayo, K.E.
Mol. Endocrinol. 5, 967-978, 1991
A;Title: Transient expression of progesterone receptor messenger RNA in ovarian granulos
A;Reference number: A23733; MUID:92049379; PMID:1840636
A;Accession: B23733
A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A;Molecule type: mRNA
A;Residues: 555-589,'W',591-624;742-910 <PAR2>
A;Cross-references: GB:S64044
R;Kraus, W.L.; Montano, M.M.; Katzenellenbogen, B.S.
Mol. Endocrinol. 7, 1603-1616, 1993
A;Title: Cloning of the rat progesterone receptor gene 5'-region and identification of t
A;Reference number: A49574; MUID:94195318; PMID:8145766
A;Accession: A49574
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-16 <KRA>
A;Cross-references: GB:S69361; NID:g546178
C;Superfamily: progesterone receptor; erbA transforming protein homology
C;Keywords: DNA binding; nucleus; steroid hormone receptor; zinc finger
F;555-819/Domain: erbA transforming protein homology <ERBA>
F;557-577/Region: zinc finger
F;593-617/Region: zinc finger

  Query Match           55.4%;  Score 759;  DB 2;  Length 923;
  Best Local Similarity  55.5%;  Pred. No. 7.1e-63;
  Matches  137;  Conservative   55;  Mismatches   55;  Indels    0;  Gaps    0;

Qy          12 PIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFRNLHV 71
               |  :|:| :||| || ||||| :||: ::||:|||:|||||: ||||:|:||||||||:
Db         675 PPLINLLMSIEPDVVYAGHDNTKPDTSSSLLTSLNQLGERQLLSVVKWSKSLPGFRNLHI 734

Qy          72 DDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHL 131
               |||: :|||||| |||| :||||: :|: :||||||||: || || :   || |: |  :
Db         735 DDQITLIQYSWMSLMVFGLGWRSYKHVSGQMLYFAPDLILNEQRMKELSFYSLCLTMWQI 794
```

```
Qy        132 SQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPT 191
              |||  ||:| :|||||| ||| : ||::||::|  |:|:| :||:|| : |  ':|
Db        795 PQEFVKLQVTHEEFLCMKVLLLLNTIPLEGLRSQSQFEEMRSSYIRELIKAIGLRQKGVV 854

Qy        192 SCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQVPKILSGK 251
              |:||||||||||:  : ::|| :  :  |:|  ::|:|||||:|:|: |:||||:|
Db        855 PSSQRFYQLTKLLDSLHDLVKQLHLYCLNTFIQSRALAVEFPEMMSEVIAAQLPKILAGM 914

Qy        252 VKPIYFH 258
              |||: ||
Db        915 VKPLLFH 921


RESULT 11
A39596
progesterone receptor B form - mouse
C;Species: Mus musculus (house mouse)
C;Date: 20-Mar-1992 #sequence_revision 20-Mar-1992 #text_change 20-Aug-1999
C;Accession: A39596; I49111
R;Schott, D.R.; Shyamala, G.; Schneider, W.; Parry, G.
Biochemistry 30, 7014-7020, 1991
A;Title: Molecular cloning, sequence analyses, and expression of complementary DNA encod
A;Reference number: A39596; MUID:91299759; PMID:2069958
A;Accession: A39596
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-923 <SCH>
A;Cross-references: GB:M68915; GB:J05333; NID:g200471; PIDN:AAA39971.1; PID:g200472
R;Hagihara, K.; Wu-Peng, X.S.; Funabashi, T.; Kato, J.; Pfaff, D.W.
Biochem. Biophys. Res. Commun. 205, 1093-1101, 1994
A;Title: Nucleic acid sequence and DNase hypersensitive sites of the 5' region of the mo
A;Reference number: I49111; MUID:95100931; PMID:7802637
A;Accession: I49111
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-9 <HAG>
A;Cross-references: EMBL:U12644; NID:g639916; PIDN:AAA66067.1; PID:g639917
C;Superfamily: progesterone receptor; erbA transforming protein homology
C;Keywords: DNA binding; nucleus; steroid hormone receptor; transcription regulation; zi
F;555-819/Domain: erbA transforming protein homology <ERBA>
F;557-577/Region: zinc finger
F;593-617/Region: zinc finger

  Query Match             55.3%;  Score 758;  DB 2;  Length 923;
  Best Local Similarity   55.1%;  Pred. No. 8.9e-63;
  Matches  136;  Conservative   56;  Mismatches   55;  Indels    0;  Gaps    0;

Qy         12 PIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFRNLHV 71
              |  :|:| :||| |: ||||| :||: ::||:|||:||||||: ||||:|:||||||||:
Db        675 PPLINLLMSIEPDVIYAGHDNTKPDTSSSLLTSLNQLGERQLLSVVKWSKSLPGFRNLHI 734

Qy         72 DDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHL 131
              |||: :|||||| |||| :|||: :|: :|||||||: || || :   || |: |  :
Db        735 DDQITLIQYSWMSLMVFGLGWRSYKHVSGQMLYFAPDLILNEQRMKELSFYSLCLTMWQI 794

Qy        132 SQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPT 191
              |||  ||:| :|||||| ||| : ||::||::|  |:|:| :||:|| : |  ::|
Db        795 PQEFVKLQVTHEEFLCMKVLLLLNTIPLEGLRSQSQFEEMRSSYIRELIKAIGLRQKGVV 854

Qy        192 SCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQVPKILSGK 251
              |:||||||||||:  : ::|| :  :  |:|  ::|:|||||:|:|: |:||||:|
Db        855 PTSQRFYQLTKLLDSLHDLVKQLHLYCLNTFIQSRTLAVEFPEMMSEVIAAQLPKILAGM 914

Qy        252 VKPIYFH 258
              |||: ||
Db        915 VKPLLFH 921


RESULT 12
A35466
progesterone receptor form B - chicken
N;Contains: progesterone receptor form A
C;Species: Gallus gallus (chicken)
C;Date: 18-Nov-1994 #sequence_revision 18-Nov-1994 #text_change 20-Aug-1999
C;Accession: A35466; S06284; A40903; A24661; A24312; A40911; A61552
R;Jeltsch, J.M.; Turcotte, B.; Garnier, J.M.; Lerouge, T.; Krozowski, Z.; Gronemeyer, H.
J. Biol. Chem. 265, 3967-3974, 1990
A;Title: Characterization of multiple mRNAs originating from the chicken progesterone re
A;Reference number: A35466; MUID:90154085; PMID:2303488
A;Accession: A35466
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-786 <JEL>
A;Cross-references: GB:M32732; GB:J05240; NID:g212554; PIDN:AAA49011.1; PID:g212558
R;Gronemeyer, H.; Turcotte, B.; Quirin-Stricker, C.; Bocquel, M.T.; Meyer, M.E.; Krozows
EMBO J. 6, 3985-3994, 1987
A;Title: The chicken progesterone receptor: sequence, expression and functional analysis
A;Reference number: S06284; MUID:88166640; PMID:3443098
A;Accession: S06284
A;Molecule type: DNA
A;Residues: 1-786 <GRO>
A;Cross-references: EMBL:Y00092; NID:g63744; PIDN:CAA68282.1; PID:g63745
R;Conneely, O.M.; Dobson, A.D.W.; Tsai, M.J.; Beattie, W.G.; Toft, D.O.; Huckaby, C.S.;
Mol. Endocrinol. 1, 517-525, 1987
A;Title: Sequence and expression of a functional chicken progesterone receptor.
A;Reference number: A40903; MUID:91042592; PMID:3153474
A;Accession: A40903
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-64,'E',65-786 <CON>
A;Cross-references: GB:M37518
R;Conneely, O.M.; Sullivan, W.P.; Toft, D.O.; Birnbaumer, M.; Cook, R.G.; Maxwell, B.L.;
Science 233, 767-770, 1986
A;Title: Molecular cloning of the chicken progesterone receptor.
A;Reference number: A24661; MUID:86289413; PMID:2426779
A;Accession: A24661
A;Molecule type: mRNA
A;Residues: 128-133,'E',135-147,'E',149-164 <CO2>
A;Note: the authors translated the codon CAG for residue 7 as Glu and CAG for residue 21
R;Jeltsch, J.M.; Krozowski, Z.; Quirin-Stricker, C.; Gronemeyer, H.; Simpson, R.J.; Garn
Proc. Natl. Acad. Sci. U.S.A. 83, 5424-5428, 1986
A;Title: Cloning of the chicken progesterone receptor.
A;Reference number: A24312; MUID:86287271; PMID:2426697
A;Accession: A24312
A;Molecule type: mRNA
A;Residues: 417-490 <JE2>
A;Cross-references: GB:M14280; NID:g212607; PIDN:AAA49039.1; PID:g212608
A;Note: amino acid and corresponding nucleotide sequences are also shown for three small
R;Birnbaumer, M.; Hinrichs-Rosello, M.V.; Cook, R.G.; Schrader, W.T.; O'Malley, B.W.
Mol. Endocrinol. 1, 249-259, 1987
A;Title: Chemical and antigenic properties of pure 108,000 molecular weight chick proges
A;Reference number: A40911; MUID:88288199; PMID:3453892
A;Accession: A40911
A;Status: preliminary
A;Molecule type: protein
A;Residues: 128-133,'E',135-147,'E',149-164;546-558 <BIR>
R;Simpson, R.J.; Grego, B.; Govindan, M.V.; Gronemeyer, H.
Mol. Cell. Endocrinol. 52, 177-184, 1987
A;Title: Peptide sequencing of the chick oviduct progesterone receptor form B.
A;Reference number: A61552; MUID:88005426; PMID:3653503
A;Accession: A61552
A;Molecule type: protein
A;Residues: 136-153;168-174;195-228;526-537,'X',539;546-563 <SIM>
C;Genetics:
A;Introns: 400/2; 451/1; 490/1; 591/1; 639/2; 683/1; 735/3
C;Superfamily: progesterone receptor; erbA transforming protein homology
C;Keywords: DNA binding; nucleus; steroid hormone receptor; zinc finger
F;1-786/Product: progesterone receptor form B #status predicted <MA1>
F;128-786/Product: progesterone receptor form A #status predicted <MA2>
F;419-682/Domain: erbA transforming protein homology <ERBA>
F;421-441/Region: zinc finger
F;457-481/Region: zinc finger

  Query Match             54.9%;  Score 752;  DB 2;  Length 786;
  Best Local Similarity   56.3%;  Pred. No. 2.6e-62;
  Matches  139;  Conservative   53;  Mismatches   55;  Indels    0;  Gaps    0;
```

```
Qy          12 PIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFRNLHV 71
               |  ::||  ||| || ||:|| :|:: ::||:||| | |||| |||||:| ||||||||:
Db         538 PPMISVLRGIEPEVVYAGYDNTKPETPSSLLTSLNHLCERQLLCVVKWSKLLPGFRNLHI 597

Qy          72 DDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHL 131
               ||| :|||||| |||||||||| :|: :||||||| || || :|  || |: |  |
Db         598 DDQITLIQYSWMSLMVFAMGWRSYKHVSGQMLYFAPDLILNEQRMKESSFYSLCLSMWQL 657

Qy         132 SQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPT 191
               |||  ||:: :|||||||||| : ||::||::|  |||:| :||:|| : |  ::|
Db         658 PQEFVRLQVSQEEFLCMKALLLLNTIPLEGLRSQSQFDEMRTSYIRELVKAIGLRQKGVV 717

Qy         192 SCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQVPKILSGK 251
               : |:|||||||:||:  : ::|| |  :  ::|  :||:|||||:|:|: |:||||:|
Db         718 ANSQRFYQLTKLMDSMHDLVKQLHLFCLNTFLQSRALSVEFPEMMSEVIAAQLPKILAGM 777

Qy         252 VKPIYFH 258
               |||: ||
Db         778 VKPLLFH 784


RESULT 13
A29513
mineralocorticoid receptor - human
N;Alternate names: aldosterone receptor
C;Species: Homo sapiens (man)
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 20-Sep-1999
C;Accession: A29513
R;Arriza, J.L.; Weinberger, C.; Cerelli, G.; Glaser, T.M.; Handelin, B.L.; Housman, D.E.
Science 237, 268-275, 1987
A;Title: Cloning of human mineralocorticoid receptor complementary DNA: structural and f
A;Reference number: A29513; MUID:87263386; PMID:3037703
A;Accession: A29513
A;Molecule type: mRNA
A;Residues: 1-984 <ARR>
A;Cross-references: GB:M16801; NID:g187460; PIDN:AAA59571.1; PID:g307166
C;Genetics:
A;Gene: GDB:MLR
A;Cross-references: GDB:120188; OMIM:264350
A;Map position: 4q31-4q31
C;Superfamily: unassigned erbA-related proteins; erbA transforming protein homology
C;Keywords: DNA binding; transcription regulation; zinc finger
F;601-880/Domain: erbA transforming protein homology <ERBA>
F;603-623/Region: zinc finger
F;639-663/Region: zinc finger

  Query Match             48.5%;  Score 665;  DB 2;  Length 984;
  Best Local Similarity   52.9%;  Pred. No. 5.2e-54;
  Matches  128;  Conservative   44;  Mismatches   70;  Indels    0;  Gaps    0;

Qy          17 VLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFRNLHVDDQMA 76
               ||| ||| :| ||:|:::||:   |||:|| |  :|:: |||||| ||||:|| ::||:
Db         741 VLENIEPEIVYAGYDSSKPDTAENLLSTLNRLAGKQMIQVVKWAKVLPGFKNLPLEDQIT 800

Qy          77 VIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHLSQEFG 136
               :|||||  |  ||: |||: : ||: ||||||||||| :||:| ||  |  |  :| :|
Db         801 LIQYSWMCLSSFALSWRSYKHTNSQFLYFAPDLVFNEEKMHQSAMYELCQGMHQISLQFV 860

Qy         137 WLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPTSCSRR 196
               ||:| :|:  || ||| | || ||||:|  |:|:| ||||||  :::     |     :|
Db         861 RLQLTFEEYTIMKVLLLLSTIPKDGLKSQAAFEEMRTNYIKELRKMVTKCPNNSGQSWQR 920

Qy         197 FYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQVPKILSGKVKPIY 256
               ||||||||||:  :  :| :| |    :|| : |:|| |: |||| |:||: ||  ||:|
Db         921 FYQLTKLLDSMHDLVSDLLEFCFYTFRESHALKVEFPAMLVEIISDQLPKVESGNAKPLY 980

Qy         257 FH 258
               ||
Db         981 FH 982


RESULT 14
QRHUGA
glucocorticoid receptor, alpha splice form - human
N;Alternate names: hGR
C;Species: Homo sapiens (man)
C;Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 21-Jul-2000
C;Accession: A93370; A39837; A39779; I56596; A03246; A32196
R;Hollenberg, S.M.; Weinberger, C.; Ong, E.S.; Cerelli, G.; Oro, A.; Lebo, R.; Thompson,
Nature 318, 635-641, 1985
A;Title: Primary structure and expression of a functional human glucocorticoid receptor
A;Reference number: A93370; MUID:86092206; PMID:2867473
A;Accession: A93370
A;Molecule type: mRNA
A;Residues: 1-777 <HOL>
A;Cross-references: EMBL:X03225; GB:M10901; NID:g31679; PIDN:CAA26976.1; PID:g31680
R;Leclerc, S.; Xie, B.; Roy, R.; Govindan, M.V.
J. Biol. Chem. 266, 8711-8719, 1991
A;Title: Purification of a human glucocorticoid receptor gene promoter-binding protein.
A;Reference number: A39837; MUID:91224961; PMID:2026589
A;Accession: A39837
A;Molecule type: DNA
A;Residues: 1-394 <LEC>
A;Cross-references: GB:M69104; NID:g183605; PIDN:AAA88049.1; PID:g553322
R;Encio, I.J.; Detera-Wadleigh, S.D.
J. Biol. Chem. 266, 7182-7188, 1991
A;Title: The genomic structure of the human glucocorticoid receptor.
A;Reference number: A39779; MUID:91201378; PMID:1707881
A;Accession: A39779
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 385-404;441-459;480-498;573-593;621-640;665-685;718-737 <ENC>
A;Cross-references: GB:M60597
A;Experimental source: placenta
R;Weinberger, C.; Hollenberg, S.M.; Rosenfeld, M.G.; Evans, R.M.
Nature 318, 670-672, 1985
A;Title: Domain structure of human glucocorticoid receptor and its relationship to the v-
A;Reference number: A93373; MUID:86092211; PMID:3841189
A;Contents: annotation; domains
R;Dahlman, K.; Stroemstedt, P.E.; Rae, C.; Joernvall, H.; Flock, J.I.; Carlstedt-Duke, J.
J. Biol. Chem. 264, 804-809, 1989
A;Title: High level expression in Escherichia coli of the DNA-binding domain of the gluc
A;Reference number: A32196; MUID:89093147; PMID:2642905
A;Contents: annotation; domains
A;Note: engineered sequence expressed in Escherichia coli
R;Govindan, M.V.; Pothier, F.; Leclerc, S.; Palaniswami, R.; Xie, B.
J. Steroid Biochem. Mol. Biol. 40, 317-323, 1991
A;Title: Human glucocorticoid receptor gene promotor-homologous down regulation.
A;Reference number: I56596; MUID:92068829; PMID:1958537
A;Accession: I56596
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-394 <RES>
A;Cross-references: GB:S68378; NID:g239757; PIDN:AAB20466.1; PID:g239758
C;Comment: Alpha and beta (see PIR:QRUGB) splice forms differ in their carboxyl-terminal
C;Genetics:
A;Gene: GDB:GRL
A;Cross-references: GDB:120017; OMIM:138040
A;Map position: 5q31-5q31
A;Introns: 395/2; 451/1; 490/1; 583/1; 631/2; 675/1; 727/3
A;Note: the first intron occurs before the initiator codon
C;Superfamily: glucocorticoid receptor; erbA transforming protein homology
C;Keywords: alternative splicing; DNA binding; nucleus; steroid hormone receptor; transci
F;419-674/Domain: erbA transforming protein homology <ERBA>
F;421-441/Region: zinc finger CCCC motif
F;457-481/Region: zinc finger CCCC motif
F;728-777/Domain: steroid binding #status predicted <STB>

  Query Match             48.3%;  Score 662.5;  DB 1;  Length 777;
  Best Local Similarity   50.6%;  Pred. No. 6.5e-54;
  Matches  125;  Conservative   49;  Mismatches   72;  Indels    1;  Gaps    1;

Qy          12 PIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFRNLHV 71
               |  :::|| ||| |: ||:|:: |||   ::::|| || ||::  ||||||:|||||||:
Db         530 PTLVSLLEVIEPEVLYAGYDSSVPDSTWRIMTTLNMLGGRQVIAAVKWAKAIPGFRNLHL 589
```

```
Qy          72 DDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHL 131
               |||| ::||||| || ||:||||:   :: :| |||||: || ||    || ||  | ::
Db         590 DDQMTLLQYSWMFLMAFALGWRSYRQSSANLLCFAPDLIINEQRMTLPCMYDQCKHMLYV 649

Qy         132 SQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPT 191
               | |   ||:: :|:|||| ||| | :| ||||:|: |||:|| ||||| : |  :  | :
Db         650 SSELHRLQVSYEEYLCMKTLLLLSSVPKDGLKSQELFDEIRMTYIKELGKAIVKREGNSS 709

Qy         192 SCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQVPKILSGK 251
                 :||||||||||:  :   |  : |   :   | |::||||:||||: |:||  :|
Db         710 QNWQRFYQLTKLLDSMHEVVENLLNYCFQTFLDKTM-SIEFPEMLAEIITNQIPKYSNGN 768

Qy         252 VKPIYFH 258
               :| : ||
Db         769 IKKLLFH 775


RESULT 15
A41401
mineralocorticoid receptor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change 20-Sep-1999
C;Accession: A41401
R;Patel, P.D.; Sherman, T.G.; Goldman, D.J.; Watson, S.J.
Mol. Endocrinol. 3, 1877-1885, 1989
A;Title: Molecular cloning of a mineralocorticoid (type I) receptor complementary DNA fr
A;Reference number: A41401; MUID:90114194; PMID:2558305
A;Accession: A41401
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-981 <PAT>
A;Cross-references: GB:M36074; NID:g205340; PIDN:AAA41583.1; PID:g205341
C;Superfamily: unassigned erbA-related proteins; erbA transforming protein homology
C;Keywords: DNA binding; transcription regulation; zinc finger
F;602-877/Domain: erbA transforming protein homology <ERBA>
F;604-624/Region: zinc finger
F;640-664/Region: zinc finger

  Query Match             48.3%;  Score 662;  DB 2;  Length 981;
  Best Local Similarity   52.2%;  Pred. No. 9.9e-54;
  Matches  128;  Conservative   47;  Mismatches   64;  Indels    6;  Gaps    2;

Qy          17 VLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFRNLHVDDQMA 76
               :|| |||  | ||:||::||:  :|||:|| |  :|:: |||||| ||||:|| ::||:
Db         738 ILENIEPETVYAGYDNSKPDTAESLLSTLNRLAAKQMIQVVKWAKVLPGFKNLPLEDQIT 797

Qy          77 VIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHLSQEFG 136
               :|||||  |  ||: |||: : ||::||||||||||| :||:| ||  |  || :| :|
Db         798 LIQYSWMCLSSFALSWRSYKHTNSQLLYFAPDLVFNEEKMHQSAMYELCQGMRQISLQFV 857

Qy         137 WLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPTSCS-- 194
                ||:| :|:  || ||| |-:| ||||:|  |:|:| ||||| :::    | | |
Db         858 RLQLTFEEYSIMKVLLLLSTVPKDGLKSQAAFEEMRTNYIKELRKMVT---KCPNSSGQS 914

Qy         195 -RRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQVPKILSGKVK 253
              .  :|||||||||||:  :  :| :| |    :|  : |:|| |: |||: |:||: ||  |
Db         915 WQRFYQLTKLLDSMHDLVSDLLEFCFYTFRESQALKVEFPAMLVEIITDQLPKVESGNAK 974

Qy         254 PIYFH 258
               |:|||
Db         975 PLYFH 979


Search completed: October  9, 2003, 11:43:05
Job time : 20 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 9, 2003, 11:41:18 ; Search time 27 Seconds
(without alignments)
1551.607 Million cell updates/sec

Title: US-09-687-609A-1
Perfect score: 1371
Sequence: 1 GSHMIEGYECQPIFLNVLEA.....SVQVPKILSGKVKPIYFHTQ 260

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 600653 seqs, 161128416 residues

Total number of hits satisfying chosen parameters: 600653

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1349.5	98.4	263	12	US-10-220-981-1	Sequence 1, Appli
2	1349.5	98.4	388	10	US-09-997-267-2	Sequence 2, Appli
3	1349.5	98.4	907	14	US-10-008-739A-2	Sequence 2, Appli
4	1349.5	98.4	919	15	US-10-205-823-36	Sequence 36, Appl
5	1243	90.7	246	10	US-09-885-827-1	Sequence 1, Appli
6	813.5	59.3	294	10	US-09-997-267-4	Sequence 4, Appli
7	765	55.8	258	12	US-10-220-981-3	Sequence 3, Appli
8	765	55.8	314	10	US-09-887-280-4	Sequence 4, Appli
9	765	55.8	933	15	US-10-207-655-63	Sequence 63, Appl
10	728	53.1	240	10	US-09-905-176-13	Sequence 13, Appl
11	666	48.6	215	10	US-09-905-176-11	Sequence 11, Appl
12	662.5	48.3	777	15	US-10-153-668-226	Sequence 226, App
13	662	48.3	689	15	US-10-202-846-2	Sequence 2, Appli
14	639.5	46.6	1070	15	US-10-001-486B-2	Sequence 2, Appli
15	630.5	46.0	284	10	US-09-853-450-20	Sequence 20, Appl
16	224	16.3	1099	15	US-10-259-864-4	Sequence 4, Appli
17	218	15.9	49	15	US-10-153-398-12	Sequence 12, Appl
18	217	15.8	595	12	US-10-437-107-30	Sequence 30, Appl
19	217	15.8	595	15	US-10-052-092-30	Sequence 30, Appl
20	216.5	15.8	595	12	US-10-437-107-31	Sequence 31, Appl
21	216.5	15.8	595	15	US-10-052-092-31	Sequence 31, Appl
22	215	15.7	575	10	US-09-893-666A-2	Sequence 2, Appli
23	213.5	15.6	323	9	US-09-737-255-5	Sequence 5, Appli
24	213.5	15.6	910	10	US-09-908-153B-40	Sequence 40, Appl
25	213.5	15.6	922	10	US-09-908-153B-42	Sequence 42, Appl
26	211.5	15.4	596	15	US-10-278-481-16	Sequence 16, Appl
27	209	15.2	460	15	US-10-295-370-4	Sequence 4, Appli
28	209	15.2	460	15	US-10-292-356-4	Sequence 4, Appli
29	209	15.2	460	16	US-10-292-324-4	Sequence 4, Appli
30	208.5	15.2	660	10	US-09-853-033-4	Sequence 4, Appli
31	207.5	15.1	414	12	US-10-157-899A-2	Sequence 2, Appli
32	207.5	15.1	414	12	US-10-157-899A-6	Sequence 6, Appli
33	207.5	15.1	414	12	US-10-157-899A-10	Sequence 10, Appl
34	207.5	15.1	438	12	US-10-157-899A-4	Sequence 4, Appli
35	207.5	15.1	438	12	US-10-157-899A-12	Sequence 12, Appl
36	207.5	15.1	438	12	US-10-157-899A-16	Sequence 16, Appl
37	207.5	15.1	472	10	US-09-965-703-61	Sequence 61, Appl
38	207.5	15.1	595	10	US-09-853-033-2	Sequence 2, Appli
39	207.5	15.1	660	10	US-09-853-033-8	Sequence 8, Appli
40	206.5	15.1	241	10	US-09-905-176-14	Sequence 14, Appl
41	206.5	15.1	484	15	US-10-278-481-13	Sequence 13, Appl
42	206.5	15.1	485	15	US-10-278-481-2	Sequence 2, Appli
43	206	15.0	511	12	US-10-006-760-19	Sequence 19, Appl
44	204.5	14.9	243	9	US-09-903-876-1	Sequence 1, Appli
45	204.5	14.9	414	12	US-10-157-899A-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-10-220-981-1
; Sequence 1, Application US/10220981
; Publication No. US20030167999A1
; GENERAL INFORMATION:
; APPLICANT: Schering AG
; APPLICANT: Donner, Peter
; APPLICANT: Egner, Ursula
; APPLICANT: Carrondo, Maria A
; APPLICANT: Matias, Pedro M
; TITLE OF INVENTION: Crystal
; FILE REFERENCE: P008463WO CTH
; CURRENT APPLICATION NUMBER: US/10/220,981
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: GB 0005689.5
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-220-981-1

Query Match 98.4%; Score 1349.5; DB 12; Length 263;
Best Local Similarity 99.6%; Pred. No. 1.1e-142;
Matches 258; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 64

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 124

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        125 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 184

Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 244

Qy        242 VQVPKILSGKVKPIYFHTQ 260
              ||||||||||||||||||
Db        245 VQVPKILSGKVKPIYFHTQ 263


RESULT 2
US-09-997-267-2
; Sequence 2, Application US/09997267
; Patent No. US20020165381A1
; GENERAL INFORMATION:
;  APPLICANT: AHRENS-FATH, ISABELLE
;  APPLICANT:  HAENDLER, BERNARD
;  TITLE OF INVENTION: HUMAN ANDROGEN RECEPTOR VARIANTS
;  FILE REFERENCE: SCH-1793
;  CURRENT APPLICATION NUMBER: US/09/997,267
;  CURRENT FILING DATE:  2001-11-30
;  PRIOR APPLICATION NUMBER: 60/255,078
;  PRIOR FILING DATE: 2000-12-14
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 388
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-997-267-2

  Query Match             98.4%;  Score 1349.5;  DB 10;  Length 388;
  Best Local Similarity   99.6%;  Pred. No. 2e-142;
  Matches  258;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 189

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 249

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 309

Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 369

Qy        242 VQVPKILSGKVKPIYFHTQ 260
              ||||||||||||||||||
Db        370 VQVPKILSGKVKPIYFHTQ 388


RESULT 3
US-10-008-739A-2
; Sequence 2, Application US/10008739A
; Publication No. US20020161194A1
; GENERAL INFORMATION:
;  APPLICANT: Pfizer Inc.
;  APPLICANT:  Castleberry, Tessa A.
;  APPLICANT:  Lu, Bihong
;  APPLICANT:  Owen, Thomas A.
;  APPLICANT:  Smock, Steven L.
;  TITLE OF INVENTION: The Canine Androgen Receptor
;  FILE REFERENCE: PC10893AGPR
;  CURRENT APPLICATION NUMBER: US/10/008,739A
;  CURRENT FILING DATE:  2002-04-15
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 907
;   TYPE: PRT
;   ORGANISM: Canine
US-10-008-739A-2

  Query Match             98.4%;  Score 1349.5;  DB 14;  Length 907;
  Best Local Similarity   99.6%;  Pred. No. 6.9e-142;
  Matches  258;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        650 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 708

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        709 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 768

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        769 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 828

Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        829 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 888

Qy        242 VQVPKILSGKVKPIYFHTQ 260
              ||||||||||||||||||
Db        889 VQVPKILSGKVKPIYFHTQ 907


RESULT 4
US-10-205-823-36
; Sequence 36, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
;  APPLICANT: Schlegel, Robert
;  APPLICANT:  Monahan, John E.
;  APPLICANT:  Endege, Wilson O.
;  APPLICANT:  Gannavarapu, Manjula
;  APPLICANT:  Gorbatcheva, Bella
;  APPLICANT:  Hoersch, Sebastian
;  APPLICANT:  Kamatkar, Shubhangi
;  APPLICANT:  Wonsey, Angela M.
;  APPLICANT:  Glatt, Karen
;  APPLICANT:  Zhao, Xumei
;  APPLICANT:  Anderson, Dustin
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
;  TITLE OF INVENTION:  METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
;  TITLE OF INVENTION:  THERAPY OF PROSTATE CANCER
;  FILE REFERENCE: MRI-044
;  CURRENT APPLICATION NUMBER: US/10/205,823
;  CURRENT FILING DATE:  2002-07-25
;  PRIOR APPLICATION NUMBER: 60/307,982
;  PRIOR FILING DATE: 2001-07-25
;  PRIOR APPLICATION NUMBER: 60/314,356
;  PRIOR FILING DATE: 2001-08-22
;  PRIOR APPLICATION NUMBER: 60/325,020
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: 60/341,746
;  PRIOR FILING DATE: 2001-12-12
;  PRIOR APPLICATION NUMBER: 60/362,158
;  PRIOR FILING DATE: 2002-03-05
;  NUMBER OF SEQ ID NOS: 455
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 36
;   LENGTH: 919
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-205-823-36

  Query Match             98.4%;  Score 1349.5;  DB 15;  Length 919;
  Best Local Similarity   99.6%;  Pred. No. 7e-142;
```

```
  Matches  258;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 720

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 780

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 840

Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 900

Qy        242 VQVPKILSGKVKPIYFHTQ 260
              |||||||||||||||||||
Db        901 VQVPKILSGKVKPIYFHTQ 919


RESULT 5
US-09-885-827-1
; Sequence 1, Application US/09885827
; Patent No. US20020173445A1
; GENERAL INFORMATION:
;  APPLICANT: Salvati, Mark
;  APPLICANT:  Attar, Ricardo M
;  APPLICANT:  Gottardis, Marco M
;  APPLICANT:  Krystek Jr., Stanley R
;  APPLICANT:  Sack, John S
;  TITLE OF INVENTION: SELECTIVE ANDROGEN RECEPTOR MODULATORS AND METHODS FOR
;  TITLE OF INVENTION:  THEIR IDENTIFICATION, DESIGN AND USE
;  FILE REFERENCE: LD0250(NP)
;  CURRENT APPLICATION NUMBER: US/09/885,827
;  CURRENT FILING DATE:  2001-06-20
;  PRIOR APPLICATION NUMBER: 60/284,438
;  PRIOR FILING DATE: 2001-04-18
;  PRIOR APPLICATION NUMBER: 60/214,392
;  PRIOR FILING DATE: 2000-06-28
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 246
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-885-827-1

  Query Match             90.7%;  Score 1243;  DB 10;  Length 246;
  Best Local Similarity   97.2%;  Pred. No. 8.9e-131;
  Matches  239;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy         13 IFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFRNLHVD 72
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db          1 IFLNVLEAIEPGVVCAGHDNAQPDSFAALLSSLNELGERQLVHVVKWAKALPGFRNLHVD 60

Qy         73 DQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHLS 132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHLS 120

Qy        133 QEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPTS 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||       |
Db        121 QEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACAAAAAAS 180

Qy        193 CSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQVPKILSGKV 252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQVPKILSGKV 240

Qy        253 KPIYFH 258
              ||||||
Db        241 KPIYFH 246


RESULT 6
US-09-997-267-4
; Sequence 4, Application US/09997267
; Patent No. US20020165381A1
; GENERAL INFORMATION:
;  APPLICANT: AHRENS-FATH, ISABELLE
;  APPLICANT:  HAENDLER, BERNARD
;  TITLE OF INVENTION: HUMAN ANDROGEN RECEPTOR VARIANTS
;  FILE REFERENCE: SCH-1793
;  CURRENT APPLICATION NUMBER: US/09/997,267
;  CURRENT FILING DATE:  2001-11-30
;  PRIOR APPLICATION NUMBER: 60/255,078
;  PRIOR FILING DATE: 2000-12-14
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 294
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-997-267-4

  Query Match             59.3%;  Score 813.5;  DB 10;  Length 294;
  Best Local Similarity   99.4%;  Pred. No. 1.5e-82;
  Matches  154;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 189

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 249

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSI 156
              ||||||||||||||||||||||||||||||||||
Db        250 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSI 284


RESULT 7
US-10-220-981-3
; Sequence 3, Application US/10220981
; Publication No. US20030167999A1
; GENERAL INFORMATION:
;  APPLICANT: Schering AG
;  APPLICANT:  Donner, Peter
;  APPLICANT:  Egner, Ursula
;  APPLICANT:  Carrondo, Maria A
;  APPLICANT:  Matias, Pedro M
;  TITLE OF INVENTION: Crystal
;  FILE REFERENCE: P008463WO CTH
;  CURRENT APPLICATION NUMBER: US/10/220,981
;  CURRENT FILING DATE:  2002-09-09
;  PRIOR APPLICATION NUMBER: GB 0005689.5
;  PRIOR FILING DATE: 2000-03-09
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 258
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-220-981-3

  Query Match             55.8%;  Score 765;  DB 12;  Length 258;
  Best Local Similarity   55.5%;  Pred. No. 3.3e-77;
  Matches  137;  Conservative   57;  Mismatches   53;  Indels    0;  Gaps    0;

Qy         12 PIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFRNLHV 71
              |  :|:| :||| |: ||||| :||: ::||:|||:|||||: ||||:|:|||||||:
Db         10 PPLINLLMSIEPDVIYAGHDNTKPDTSSSLLTSLNQLGERQLLSVVKWSKSLPGFRNLHI 69

Qy         72 DDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHL 131
```

```
              |||: :|||||| |||| :||||: :|: :|||||||: || || :|  || |: |  :
Db         70 DDQITLIQYSWMSLMVFGLGWRSYKHVSGQMLYFAPDLILNEQRMKESSFYSLCLTMWQI 129

Qy        132 SQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPT 191
              |||  ||:: :|||||| ||| : ||::||::|  |:|:| :||:|| : |  ::|
Db        130 PQEFVKLQVSQEEFLCMKVLLLLNTIPLEGLRSQTQFEEMRSSYIRELIKAIGLRQKGVV 189

Qy        192 SCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQVPKILSGK 251
              | |:|||||||||::  : ::|| :  :  |:|  :||:|||||:|:|: |:||||:|
Db        190 SSSQRFYQLTKLLDNLHDLVKQLHLYCLNTFIQSRALSVEFPEMMSEVIAAQLPKILAGM 249

Qy        252 VKPIYFH 258
              |||: ||
Db        250 VKPLLFH 256


RESULT 8
US-09-887-280-4
; Sequence 4, Application US/09887280
; Publication No. US20020197670A1
; GENERAL INFORMATION:
;  APPLICANT: PRICE, THOMAS M.
;  TITLE OF INVENTION: MEMBRANE ASSOCIATED PROGESTERONE RECEPTOR
;  FILE REFERENCE: GHS-338
;  CURRENT APPLICATION NUMBER: US/09/887,280
;  CURRENT FILING DATE:  2001-09-17
;  PRIOR APPLICATION NUMBER: 60/213,340
;  PRIOR FILING DATE: 2000-06-22
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 314
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-887-280-4

  Query Match             55.8%;  Score 765;  DB 10;  Length 314;
  Best Local Similarity   55.5%;  Pred. No. 4.4e-77;
  Matches  137;  Conservative   57;  Mismatches   53;  Indels    0;  Gaps    0;

Qy         12 PIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFRNLHV 71
              |  :|:| :||| |: ||||| :||: ::||:|||:||||||: ||||:|:||||||||:
Db         66 PPLINLLMSIEPDVIYAGHDNTKPDTSSSLLTSLNQLGERQLLSVVKWSKSLPGFRNLHI 125

Qy         72 DDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHL 131
              |||: :|||||| |||| :||||: :|: :|||||||: || || :|  || |: |  :
Db        126 DDQITLIQYSWMSLMVFGLGWRSYKHVSGQMLYFAPDLILNEQRMKESSFYSLCLTMWQI 185

Qy        132 SQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPT 191
              |||  ||:: :|||||| ||| : ||::||::|  |:|:| :||:|| : |  ::|
Db        186 PQEFVKLQVSQEEFLCMKVLLLLNTIPLEGLRSQTQFEEMRSSYIRELIKAIGLRQKGVV 245

Qy        192 SCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQVPKILSGK 251
              | |:|||||||||::  : ::|| :  :  |:|  :||:|||||:|:|: |:||||:|
Db        246 SSSQRFYQLTKLLDNLHDLVKQLHLYCLNTFIQSRALSVEFPEMMSEVIAAQLPKILAGM 305

Qy        252 VKPIYFH 258
              |||: ||
Db        306 VKPLLFH 312


RESULT 9
US-10-207-655-63
; Sequence 63, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
;  APPLICANT: Ledbetter, Jeffrey A.
;  APPLICANT:  Hayden-Ledbetter, Martha S.
;  TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
;  FILE REFERENCE: 390069.401C1
;  CURRENT APPLICATION NUMBER: US/10/207,655
;  CURRENT FILING DATE:  2002-07-25
;  NUMBER OF SEQ ID NOS: 426
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 63
;   LENGTH: 933
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-207-655-63

  Query Match             55.8%;  Score 765;  DB 15;  Length 933;
  Best Local Similarity   55.5%;  Pred. No. 2.1e-76;
  Matches  137;  Conservative   57;  Mismatches   53;  Indels    0;  Gaps    0;

Qy         12 PIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFRNLHV 71
              |  :|:| :||| |: ||||| :||: ::||:|||:||||||: ||||:|:||||||||:
Db        685 PPLINLLMSIEPDVIYAGHDNTKPDTSSSLLTSLNQLGERQLLSVVKWSKSLPGFRNLHI 744

Qy         72 DDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHL 131
              |||: :|||||| |||| :||||: :|: :|||||||: || || :|  || |: |  :
Db        745 DDQITLIQYSWMSLMVFGLGWRSYKHVSGQMLYFAPDLILNEQRMKESSFYSLCLTMWQI 804

Qy        132 SQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPT 191
              |||  ||:: :|||||| ||| : ||::||::|  |:|:| :||:|| : |  ::|
Db        805 PQEFVKLQVSQEEFLCMKVLLLLNTIPLEGLRSQTQFEEMRSSYIRELIKAIGLRQKGVV 864

Qy        192 SCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQVPKILSGK 251
              | |:|||||||||::  : ::|| :  :  |:|  :||:|||||:|:|: |:||||:|
Db        865 SSSQRFYQLTKLLDNLHDLVKQLHLYCLNTFIQSRALSVEFPEMMSEVIAAQLPKILAGM 924

Qy        252 VKPIYFH 258
              |||: ||
Db        925 VKPLLFH 931


RESULT 10
US-09-905-176-13
; Sequence 13, Application US/09905176
; Patent No. US20020150906A1
; GENERAL INFORMATION:
;  APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY
;  APPLICANT:  Debe, Derek A.
;  TITLE OF INVENTION: METHOD FOR DETERMINING THREE-DIMENSIONAL PROTEIN STRUCTURE FROM P
;  TITLE OF INVENTION:  PROTEIN SEQUENCE
;  FILE REFERENCE: 265/297
;  CURRENT APPLICATION NUMBER: US/09/905,176
;  CURRENT FILING DATE:  2002-04-05
;  PRIOR APPLICATION NUMBER: US 60/218,016
;  PRIOR FILING DATE: 2000-07-12
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
;   LENGTH: 240
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-905-176-13

  Query Match             53.1%;  Score 728;  DB 10;  Length 240;
  Best Local Similarity   55.3%;  Pred. No. 4.1e-73;
  Matches  131;  Conservative   55;  Mismatches   51;  Indels    0;  Gaps    0;

Qy         12 PIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFRNLHV 71
              |  :|:| :||| |: ||||| :||: ::||:|||:||||||: ||||:|:||||||||:
Db          4 PPLINLLMSIEPDVIYAGHDNTKPDTSSSLLTSLNQLGERQLLSVVKWSKSLPGFRNLHI 63

Qy         72 DDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHL 131
              |||: :|||||| |||| :||||: :|: :|||||||: || || :|  || |: |  :
Db         64 DDQITLIQYSWMSLMVFGLGWRSYKHVSGQMLYFAPDLILNEQRMKESSFYSLCLTMWQI 123

Qy        132 SQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPT 191
              |||  ||:: :|||||| ||| : ||::||::|  |:|:| :||:|| : |  ::|
Db        124 PQEFVKLQVSQEEFLCMKVLLLLNTIPLEGLRSQTQFEEMRSSYIRELIKAIGLRQKGVV 183

Qy        192 SCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQVPKIL 248
              | |:|||||||||::  : ::|| :  :  |:|  :||:|||||:|:|: |:||||
Db        184 SSSQRFYQLTKLLDNLHDLVKQLHLYCLNTFIQSRALSVEFPEMMSEVIAAQLPKIL 240
```

```
RESULT 11
US-09-905-176-11
; Sequence 11, Application US/09905176
; Patent No. US20020150906A1
; GENERAL INFORMATION:
;  APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY
;  APPLICANT:  Debe, Derek A.
;  TITLE OF INVENTION: METHOD FOR DETERMINING THREE-DIMENSIONAL PROTEIN STRUCTURE FROM P
;  TITLE OF INVENTION:  PROTEIN SEQUENCE
;  FILE REFERENCE: 265/297
;  CURRENT APPLICATION NUMBER: US/09/905,176
;  CURRENT FILING DATE:  2002-04-05
;  PRIOR APPLICATION NUMBER: US 60/218,016
;  PRIOR FILING DATE: 2000-07-12
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
;   LENGTH: 215
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-905-176-11

  Query Match            48.6%;  Score 666;  DB 10;  Length 215;
  Best Local Similarity  55.6%;  Pred. No. 3.1e-66;
  Matches  119;  Conservative   50;  Mismatches   45;  Indels    0;  Gaps    0;

Qy         28 AGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFRNLHVDDQMAVIQYSWMGLMV 87
              ||||| :||: ::||:|||:|||||||: ||||:|:|||||||:|||: :|||||| |||
Db          1 AGHDNTKPDTSSSLLTSLNQLGERQLLSVVKWSKSLPGFRNLHIDDQITLIQYSWMSLMV 60

Qy         88 FAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHLSQEFGWLQITPQEFLC 147
              | :|||: :|: :|||||||: || || :|  || |: |  : |||  ||:: :||||
Db         61 FGLGWRSYKHVSGQMLYFAPDLILNEQRMKESSFYSLCLTMWQIPQEFVKLQVSQEEFLC 120

Qy        148 MKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPTSCSRRFYQLTKLLDSV 207
              || ||| : ||::||::|  |:|:| :||:|| : |  ::|    | |:||||||||||::
Db        121 MKVLLLLNTIPLEGLRSQTQFEEMRSSYIRELIKAIGLRQKGVVSSSQRFYQLTKLLDNL 180

Qy        208 QPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              : ::|| :  :  |:|  :||:|||||:|:|:
Db        181 HDLVKQLHLYCLNTFIQSRALSVEFPEMMSEVIA 214


RESULT 12
US-10-153-668-226
; Sequence 226, Application US/10153668
; Publication No. US20030092616A1
; GENERAL INFORMATION:
;  APPLICANT: HONDA, Goichi
;  APPLICANT:  MATSUDA, Akio
;  APPLICANT:  MURAMATSU, Shuji
;  APPLICANT:  ISHIZAWA, Kenya
;  TITLE OF INVENTION: STAT6 Activating Gene
;  FILE REFERENCE: 1254-0207P
;  CURRENT APPLICATION NUMBER: US/10/153,668
;  CURRENT FILING DATE:  2002-05-24
;  PRIOR APPLICATION NUMBER: US 60/293,172
;  PRIOR FILING DATE: 2001-05-25
;  PRIOR APPLICATION NUMBER: US 60/316,031
;  PRIOR FILING DATE: 2001-08-31
;  PRIOR APPLICATION NUMBER: US 60/328,403
;  PRIOR FILING DATE: 2001-10-12
;  PRIOR APPLICATION NUMBER: JP 2001-157043
;  PRIOR FILING DATE: 2001-05-25
;  PRIOR APPLICATION NUMBER: JP 2001-260681
;  PRIOR FILING DATE: 2001-08-30
;  PRIOR APPLICATION NUMBER: JP 2001-313175
;  PRIOR FILING DATE: 2001-10-10
;  NUMBER OF SEQ ID NOS: 488
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 226
;   LENGTH: 777
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-153-668-226

  Query Match            48.3%;  Score 662.5;  DB 15;  Length 777;
  Best Local Similarity  50.6%;  Pred. No. 4.9e-65;
  Matches  125;  Conservative   49;  Mismatches   72;  Indels    1;  Gaps    1;

Qy         12 PIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFRNLHV 71
                |  :::|| ||| |: ||:|:: |||   ::::|| || ||::  ||||||:|||||||:
Db        530 PTLVSLLEVIEPEVLYAGYDSSVPDSTWRIMTTLNMLGGRQVIAAVKWAKAIPGFRNLHL 589

Qy         72 DDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHL 131
              |||| ::||||| || ||:||||:   :: :| |||||: || ||    || ||  | ::
Db        590 DDQMTLLQYSWMFLMAFALGWRSYRQSSANLLCFAPDLIINEQRMTLPCMYDQCKHMLYV 649

Qy        132 SQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPT 191
              | |   ||:: :|:|||| ||| | :| ||||:|: |||:|| ||||| : |  :  | :
Db        650 SSELHRLQVSYEEYLCMKTLLLLSSVPKDGLKSQELFDEIRMTYIKELGKAIVKREGNSS 709

Qy        192 SCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQVPKILSGK 251
              :|||||||||:  :   |  : |   :   | |::||||:||||: |:||  :|
Db        710 QNWQRFYQLTKLLDSMHEVVENLLNYCFQTFLDKTM-SIEFPEMLAEIITNQIPKYSNGN 768

Qy        252 VKPIYFH 258
              :| : ||
Db        769 IKKLLFH 775


RESULT 13
US-10-202-846-2
; Sequence 2, Application US/10202846
; Publication No. US20030083487A1
; GENERAL INFORMATION:
;  APPLICANT: NAGAHAMA, Yoshitaka
;  APPLICANT:  IKEUCHI, Toshitaka
;  APPLICANT:  KOBAYASHI, Toru
;  APPLICANT:  TODO, Takashi
;  TITLE OF INVENTION: PROGESTOGEN RECEPTOR POLYPEPTIDES, TRANSGENIC CELLS IN WHICH GENES
;  TITLE OF INVENTION:  ENCODING SAID POLYPEPTIDES ARE INTRODUCED AND METHOD FOR DETECTI
;  TITLE OF INVENTION:  FOR PROGESTOGENS BY USING SAID TRANSGENIC CELLS
;  FILE REFERENCE: 113343
;  CURRENT APPLICATION NUMBER: US/10/202,846
;  CURRENT FILING DATE:  2002-07-26
;  PRIOR APPLICATION NUMBER: JP 2001/235,725
;  PRIOR FILING DATE: 2001-08-03
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 689
;   TYPE: PRT
;   ORGANISM: Anguilla japonica
US-10-202-846-2

  Query Match            48.3%;  Score 662;  DB 15;  Length 689;
  Best Local Similarity  47.5%;  Pred. No. 4.7e-65;
  Matches  121;  Conservative   59;  Mismatches   75;  Indels    0;  Gaps    0;

Qy          4 MIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKAL 63
              |:  :  |  |::|| |||  | :|:|  ||::  ||:||| | ||||: :|:|:|:|
Db        433 MVRELQFTPQILSILENIEPETVYSGYDATQPETPHLLLNSLNGLCERQLLWIVRWSKSL 492

Qy         64 PGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYS 123
              ||||:||::||| :|||||| ||||::|||| ||    |||||||:  | :|  | :
Db        493 PGFRSLHINDQMTLIQYSWMSLMVFSLGWRSFQNVTREFLYFAPDLILGEEKMRNSPISD 552

Qy        124 QCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRII 183
              |: |: : | |  ||:| :|||||| ||| : :|::||::|  |||:|  ||:|| : |
Db        553 LCMAMQIIPQAFDNLQVTKEEFLCMKVLLLLNTVPLEGLRSQAQFDEMRHGYIRELTKAI 612

Qy        184 ACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQ 243
                  :   : |:||| ||||:|::  | |::: :     |::  : |:|||||:|:|: |
```

```
Db        613 QLTERGVMASSQRFYHLTKLMDAMHEIVRKVNLYCLSTFIQAEAMQVEFPEMMSEVITSQ 672

Qy        244 VPKILSGKVKPIYFH 258
              :||:|:| |:|: ||
Db        673 LPKVLAGMVRPLLFH 687


RESULT 14
US-10-001-486B-2
; Sequence 2, Application US/10001486B
; Publication No. US20030082642A1
;   GENERAL INFORMATION:
;        APPLICANT: The Government of the United States of America
;                   as represented by the Secretary
;                   Department of Health and Human Services
;                   Washington, D.C.
;                   Htun Ph.D., Han
;                   Hager Ph.D., Gordon L.
;        TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MONITORING
;                            DNA BINDING MOLECULES IN LIVING CELLS
;        NUMBER OF SEQUENCES: 9
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Needle & Rosenberg
;             STREET: 127 Peachtree Street, Suite 1200
;             CITY: Atlanta
;             STATE: Georgia
;             COUNTRY: USA
;             ZIP: 30303
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/001,486B
;             FILING DATE: 15-No. US20030082642A1-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 60/008,373
;             FILING DATE: 08 Dec 1995
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Selby, Elizabeth
;             REGISTRATION NUMBER: 38298
;             REFERENCE/DOCKET NUMBER: 14014.0183
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 404-688-0770
;             TELEFAX: 404-688-9880
;   INFORMATION FOR SEQ ID NO: 2:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1070 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-001-486B-2

  Query Match             46.6%;  Score 639.5;  DB 15;  Length 1070;
  Best Local Similarity   49.4%;  Pred. No. 3e-62;
  Matches  122;  Conservative   48;  Mismatches   76;  Indels    1;  Gaps    1;

Qy         12 PIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFRNLHV 71
              |  :::|| ||| |: ||:|:: |||   ::::|| || ||::  |||||: | ||||:
Db        823 PTLVSLLEVIEPEVLYAGYDSSVPDSAWRIMTTLNMLGGRQVIAAVKWAKAILGLRNLHL 882

Qy         72 DDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHL 131
              |||| ::||||| || ||:||||:   :  :| |||||: || ||    || ||  |  :
Db        883 DDQMTLLQYSWMFLMAFALGWRSYRQSSGNLLCFAPDLIINEQRMSLPGMYDQCKHMLFV 942

Qy        132 SQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPT 191
              | |   ||:: :|:|||| ||| | :| :|||:|: |||:|| ||||| : |  :  | :
Db        943 SSELQRLQVSYEEYLCMKTLLLLSSVPKEGLKSQELFDEIRMTYIKELGKAIVKREGNSS 1002

Qy        192 SCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQVPKILSGK 251
```

```
              :||||||||||:  :   |  : |   :   | |::||||:||||: |:||  :|
Db       1003 QNWQRFYQLTKLLDSMHEVVENLLTYCFQTFLDKTM-SIEFPEMLAEIITNQIPKYSNGN 1061

Qy        252 VKPIYFH 258
              :| : ||
Db       1062 IKKLLFH 1068


RESULT 15
US-09-853-450-20
; Sequence 20, Application US/09853450
; Publication No. US20020194645A1
; GENERAL INFORMATION:
;  APPLICANT: Yanofsky, Martin F.
;  APPLICANT:  Pelaz, Soraya
;  APPLICANT:  Ditta, Gary
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: Combinations of Genes for Producing Seed Plants
;  TITLE OF INVENTION:  Exhibiting Modulated Reproductive Development
;  FILE REFERENCE: 19452A-002400US
;  CURRENT APPLICATION NUMBER: US/09/853,450
;  CURRENT FILING DATE:  2001-05-09
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
;   LENGTH: 284
;   TYPE: PRT
;   ORGANISM: Rattus norvegicus
;   FEATURE:
;   OTHER INFORMATION: rat glucocorticoid receptor ligand binding domain
US-09-853-450-20

  Query Match             46.0%;  Score 630.5;  DB 10;  Length 284;
  Best Local Similarity   49.0%;  Pred. No. 4.4e-62;
  Matches  121;  Conservative   48;  Mismatches   77;  Indels    1;  Gaps    1;

Qy         12 PIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFRNLHV 71
              |  :::|| ||| |: ||:|:: |||   ::::|| || ||::  |||||: | ||||:
Db         37 PTLVSLLEVIEPEVLYAGYDSSVPDSAWRIMTTLNMLGGRQVIAAVKWAKAILGLRNLHL 96

Qy         72 DDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHL 131
              |||| ::||||| || ||:||||:   :  :| |||||: || ||    || ||  |  :
Db         97 DDQMTLLQYSWMFLMAFALGWRSYRQSSGNLLCFAPDLIINEQRMSLPCMYDQCKHMLFV 156

Qy        132 SQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPT 191
              | |   ||:: :|:|||| ||| | :| :|||:|: |||:|| ||||| : |  :  | :
Db        157 SSELQRLQVSYEEYLCMKTLLLLSSVAKEGLKSQELFDEIRMTYIKELGKAIVKREGNSS 216

Qy        192 SCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQVPKILSGK 251
              :||||||||||:  :   |  : |   :   | |::||||:||||: |:||  :|
Db        217 QNWQRFYQLTKLLDSMHEVVENLLTYCFQTFLDKTM-SIEFPEMLAEIITNQIPKYSNGN 275

Qy        252 VKPIYFH 258
              :| : ||
Db        276 IKKLLFH 282


Search completed: October  9, 2003, 11:43:39
Job time : 28 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 9, 2003, 11:39:53 ; Search time 19 Seconds
(without alignments)
578.991 Million cell updates/sec

Title: US-09-687-609A-1
Perfect score: 1371
Sequence: 1 GSHMIEGYECQPIFLNVLEA.....SVQVPKILSGKVKPIYFHTQ 260

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1349.5	98.4	452	3	US-08-764-870-16	Sequence 16, Appl
2	1349.5	98.4	452	3	US-08-980-115-16	Sequence 16, Appl
3	1349.5	98.4	918	3	US-09-041-886-11	Sequence 11, Appl
4	766	55.9	363	6	5223606-6	Patent No. 5223606
5	765	55.8	933	3	US-08-764-870-14	Sequence 14, Appl
6	765	55.8	933	3	US-08-980-115-14	Sequence 14, Appl
7	665	48.5	984	3	US-08-764-870-15	Sequence 15, Appl
8	665	48.5	984	3	US-08-980-115-15	Sequence 15, Appl
9	662.5	48.3	777	3	US-08-764-870-13	Sequence 13, Appl
10	662.5	48.3	777	3	US-08-980-115-13	Sequence 13, Appl
11	644.5	47.0	534	3	US-08-875-223-8	Sequence 8, Appli
12	639.5	46.6	1070	4	US-09-091-042A-2	Sequence 2, Appli
13	638.5	46.6	284	2	US-08-592-214A-24	Sequence 24, Appl
14	638.5	46.6	284	3	US-09-149-976-24	Sequence 24, Appl
15	638.5	46.6	795	1	US-07-716-827C-5	Sequence 5, Appli
16	630.5	46.0	284	3	US-08-659-188-20	Sequence 20, Appl
17	630.5	46.0	284	3	US-08-655-227-20	Sequence 20, Appl
18	630.5	46.0	284	3	US-08-655-241-20	Sequence 20, Appl
19	630.5	46.0	284	4	US-09-398-326-20	Sequence 20, Appl
20	626	45.7	356	6	5223606-7	Patent No. 5223606
21	224	16.3	60	5	PCT-US94-14074-1	Sequence 1, Appli
22	211.5	15.4	596	2	US-08-836-620A-16	Sequence 16, Appl
23	209	15.2	460	4	US-09-393-839-4	Sequence 4, Appli
24	207.5	15.1	595	3	US-08-764-870-12	Sequence 12, Appl
25	207.5	15.1	595	3	US-08-980-115-12	Sequence 12, Appl
26	206.5	15.1	484	2	US-08-836-620A-13	Sequence 13, Appl
27	206.5	15.1	485	2	US-08-836-620A-2	Sequence 2, Appli
28	205	15.0	228	3	US-09-249-645-2	Sequence 2, Appli
29	205	15.0	228	4	US-09-844-132B-2	Sequence 2, Appli
30	204.5	14.9	591	2	US-08-836-620A-17	Sequence 17, Appl
31	204.5	14.9	595	3	US-09-041-886-35	Sequence 35, Appl
32	204.5	14.9	595	4	US-08-453-998-2	Sequence 2, Appli
33	204.5	14.9	651	3	US-08-693-940-3	Sequence 3, Appli
34	204.5	14.9	651	4	US-09-566-660-3	Sequence 3, Appli
35	204.5	14.9	773	3	US-08-564-264-1	Sequence 1, Appli
36	203.5	14.8	485	2	US-08-836-620A-3	Sequence 3, Appli
37	203.5	14.8	548	3	US-09-139-617-1	Sequence 1, Appli
38	203.5	14.8	548	4	US-09-561-741A-1	Sequence 1, Appli
39	203.5	14.8	548	4	US-09-558-795-1	Sequence 1, Appli
40	203	14.8	384	2	US-08-836-620A-15	Sequence 15, Appl
41	202.5	14.8	264	4	US-09-660-979-1	Sequence 1, Appli
42	200.5	14.6	229	3	US-09-249-645-1	Sequence 1, Appli
43	198.5	14.5	484	2	US-08-836-620A-14	Sequence 14, Appl
44	198.5	14.5	485	2	US-08-836-620A-5	Sequence 5, Appli
45	196.5	14.3	229	4	US-09-844-132B-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-764-870-16
; Sequence 16, Application US/08764870
; Patent No. 6236946
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Thomas S
; APPLICANT: Baxter, John D
; APPLICANT: Fletterick, Robert J
; APPLICANT: Wagner, Richard L
; APPLICANT: Kushner, Peter J
; APPLICANT: Apriletti, James W
; APPLICANT: West, Brian
; TITLE OF INVENTION: Nuclear Receptor Ligands and Ligand
; TITLE OF INVENTION: Binding Domains
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooley Godward
; STREET: Five Palo Alto Square, 3000 El Camino Real
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/764,870
; FILING DATE: 13-DEC-1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/008,540
; FILING DATE: 13-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/008,543
; FILING DATE: 13-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/008,606
; FILING DATE: 14-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Nakamura, Jackie N
; REGISTRATION NUMBER: 35,966
; REFERENCE/DOCKET NUMBER: UCAL-246/01US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650)843-5000
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 452 amino acids
; TYPE: amino acid

```
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-764-870-16

  Query Match            98.4%;  Score 1349.5;  DB 3;  Length 452;
  Best Local Similarity   99.6%;  Pred. No. 9.1e-147;
  Matches  258;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 253

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 313

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 373

Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 433

Qy        242 VQVPKILSGKVKPIYFHTQ 260
              ||||||||||||||||||
Db        434 VQVPKILSGKVKPIYFHTQ 452


RESULT 2
US-08-980-115-16
; Sequence 16, Application US/08980115
; Patent No. 6266622
; GENERAL INFORMATION:
;  APPLICANT: Scanlan, Thomas S.
;  APPLICANT: Baxter, John D.
;  APPLICANT: Fletterick, Robert J.
;  APPLICANT: Wagner, Richard L.
;  APPLICANT: Kushner, Peter J.
;  APPLICANT: Apriletti, James W.
;  APPLICANT: West, Brian L.
;  APPLICANT: Shiau, Andrew K.
;  TITLE OF INVENTION: NUCLEAR RECEPTOR LIGANDS AND LIGAND BINDING DOMAINS
;  FILE REFERENCE: UCAL-246/02US
;  CURRENT APPLICATION NUMBER: US/08/980,115
;  CURRENT FILING DATE: 1997-11-26
;  EARLIER APPLICATION NUMBER: 08/764,870
;  EARLIER FILING DATE: 1996-12-13
;  EARLIER APPLICATION NUMBER: 60/008,606
;  EARLIER FILING DATE: 1995-12-14
;  EARLIER APPLICATION NUMBER: 60/008,543
;  EARLIER FILING DATE: 1995-12-13
;  EARLIER APPLICATION NUMBER: 60/008,540
;  EARLIER FILING DATE: 1995-12-13
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
;   LENGTH: 452
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: DOMAIN
;   LOCATION: (184)..(437)
;   OTHER INFORMATION: minimal ligand binding domain
US-08-980-115-16

  Query Match            98.4%;  Score 1349.5;  DB 3;  Length 452;
  Best Local Similarity   99.6%;  Pred. No. 9.1e-147;
  Matches  258;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 253

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 313

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 373

Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 433

Qy        242 VQVPKILSGKVKPIYFHTQ 260
              ||||||||||||||||||
Db        434 VQVPKILSGKVKPIYFHTQ 452


RESULT 3
US-09-041-886-11
; Sequence 11, Application US/09041886
; Patent No. 6235872
;  GENERAL INFORMATION:
;    APPLICANT:  Bredesen, Dale E.
;    APPLICANT:  Rabizadeh, Sharroz
;    TITLE OF INVENTION:  Proapoptotic Peptides, Dependence
;    TITLE OF INVENTION:  Polypeptides and Methods of Use
;    NUMBER OF SEQUENCES:  72
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell & Flores LLP
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/041,886
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-LJ 2626
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  918 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-041-886-11

  Query Match            98.4%;  Score 1349.5;  DB 3;  Length 918;
  Best Local Similarity   99.6%;  Pred. No. 2.6e-146;
  Matches  258;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 719

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 779

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 253

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 313

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 373

Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 433

Qy        242 VQVPKILSGKVKPIYFHTQ 260
              ||||||||||||||||||
Db        434 VQVPKILSGKVKPIYFHTQ 452


RESULT 3
US-09-041-886-11
; Sequence 11, Application US/09041886
; Patent No. 6235872
;  GENERAL INFORMATION:
;    APPLICANT:  Bredesen, Dale E.
;    APPLICANT:  Rabizadeh, Sharroz
;    TITLE OF INVENTION:  Proapoptotic Peptides, Dependence
;    TITLE OF INVENTION:  Polypeptides and Methods of Use
;    NUMBER OF SEQUENCES:  72
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell & Flores LLP
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/041,886
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-LJ 2626
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  918 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-041-886-11

  Query Match            98.4%;  Score 1349.5;  DB 3;  Length 918;
  Best Local Similarity   99.6%;  Pred. No. 2.6e-146;
  Matches  258;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 719

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 779

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 839

Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 899

Qy        242 VQVPKILSGKVKPIYFHTQ 260
                |||||||||||||||||||
Db        900 VQVPKILSGKVKPIYFHTQ 918


RESULT 4
5223606-6
;Patent No. 5223606.
;    APPLICANT:  BLAUDIN DE THE, HUGHES;MARCHIO, AGNES;TIOLLAIS,
;PIERRE;DEJEAN, ANNE.
;    TITLE OF INVENTION: STEROID/THYROID HORMONE RECEPTOR-RELATED
;PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA
;    NUMBER OF SEQUENCES: 11
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/134,130
;      FILING DATE: 17-DEC-1987
;    PRIOR APPLICATION DATA:
;SEQ ID NO:6:
;      LENGTH: 363
5223606-6

  Query Match           55.9%;  Score 766;  DB 6;  Length 363;
  Best Local Similarity   55.5%;  Pred. No. 1.1e-79;
  Matches  137;  Conservative   57;  Mismatches   53;  Indels    0;  Gaps    0;

Qy         12 PIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFRNLHV 71
                |  :|:| :||| |: ||||| :||: ::||:|||:|||||: ||||:|:||||||||:
Db        115 PPLINLLMSIEPDVIYAGHDNTKPDTSSSLLTSLNQLGERQLLSVVKWSKSLPGFRNLHI 174

Qy         72 DDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHL 131
                |||: :|||||| |||| :||||: :|: :||||||||: || || :|  || |: |  :
Db        175 DDQITLIQYSWMSLMVFGLGWRSYKHVSGQMLYFAPDLILNEQRMKESSFYSLCLTMWQI 234

Qy        132 SQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPT 191
                 |||  ||:: :|||||| ||| : ||::||::|  |:|:| :||:|| : |  ::|
Db        235 PQEFVKLQVSQEEFLCMKVLLLLNTIPLEGLRSQSQFEEMRSSYIRELIKAIGLRQKGVV 294

Qy        192 SCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQVPKILSGK 251
                | |:||||||||||::  : ::|| :  :  |:|  :||:|||||:|:|: |:||||:|
Db        295 SSSQRFYQLTKLLDNLHDLVKQLHLYCLNTFIQSRALSVEFPEMMSEVIAAQLPKILAGM 354

Qy        252 VKPIYFH 258
                |||: ||
Db        355 VKPLLFH 361


RESULT 5
US-08-764-870-14
; Sequence 14, Application US/08764870
; Patent No. 6236946
;  GENERAL INFORMATION:
;    APPLICANT:  Scanlan, Thomas S
;    APPLICANT:  Baxter, John D
;    APPLICANT:  Fletterick, Robert J
;    APPLICANT:  Wagner, Richard L
;    APPLICANT:  Kushner, Peter J
;    APPLICANT:  Apriletti, James W
;    APPLICANT:  West, Brian
;    TITLE OF INVENTION:  Nuclear Receptor Ligands and Ligand
;    TITLE OF INVENTION:  Binding Domains
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooley Godward
;      STREET:  Five Palo Alto Square,  3000 El Camino Real
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/764,870
;      FILING DATE:  13-DEC-1996
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/008,540
;      FILING DATE:  13-DEC-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/008,543
;      FILING DATE:  13-DEC-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/008,606
;      FILING DATE:  14-DEC-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Nakamura, Jackie N
;      REGISTRATION NUMBER:  35,966
;      REFERENCE/DOCKET NUMBER:  UCAL-246/01US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650)843-5000
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  933 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-764-870-14

  Query Match           55.8%;  Score 765;  DB 3;  Length 933;
  Best Local Similarity   55.5%;  Pred. No. 5.6e-79;
  Matches  137;  Conservative   57;  Mismatches   53;  Indels    0;  Gaps    0;

Qy         12 PIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFRNLHV 71
                |  :|:| :||| |: ||||| :||: ::||:|||:|||||: ||||:|:||||||||:
Db        685 PPLINLLMSIEPDVIYAGHDNTKPDTSSSLLTSLNQLGERQLLSVVKWSKSLPGFRNLHI 744

Qy         72 DDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHL 131
                |||: :|||||| |||| :||||: :|: :||||||||: || || :|  || |: |  :
Db        745 DDQITLIQYSWMSLMVFGLGWRSYKHVSGQMLYFAPDLILNEQRMKESSFYSLCLTMWQI 804

Qy        132 SQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPT 191
                 |||  ||:: :|||||| ||| : ||::||::|  |:|:| :||:|| : |  ::|
Db        805 PQEFVKLQVSQEEFLCMKVLLLLNTIPLEGLRSQTQFEEMRSSYIRELIKAIGLRQKGVV 864

Qy        192 SCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQVPKILSGK 251
                | |:||||||||||::  : ::|| :  :  |:|  :||:|||||:|:|: |:||||:|
Db        865 SSSQRFYQLTKLLDNLHDLVKQLHLYCLNTFIQSRALSVEFPEMMSEVIAAQLPKILAGM 924

Qy        252 VKPIYFH 258
                |||: ||
Db        925 VKPLLFH 931


RESULT 6
US-08-980-115-14
; Sequence 14, Application US/08980115
; Patent No. 6266622
; GENERAL INFORMATION:
;  APPLICANT: Scanlan, Thomas S.
;  APPLICANT: Baxter, John D.
;  APPLICANT: Fletterick, Robert J.
;  APPLICANT: Wagner, Richard L.
;  APPLICANT: Kushner, Peter J.
;  APPLICANT: Apriletti, James W.
;  APPLICANT: West, Brian L.
;  APPLICANT: Shiau, Andrew K.
```

```
;  TITLE OF INVENTION: NUCLEAR RECEPTOR LIGANDS AND LIGAND BINDING DOMAINS
;  FILE REFERENCE: UCAL-246/02US
;  CURRENT APPLICATION NUMBER: US/08/980,115
;  CURRENT FILING DATE: 1997-11-26
;  EARLIER APPLICATION NUMBER: 08/764,870
;  EARLIER FILING DATE: 1996-12-13
;  EARLIER APPLICATION NUMBER: 60/008,606
;  EARLIER FILING DATE: 1995-12-14
;  EARLIER APPLICATION NUMBER: 60/008,543
;  EARLIER FILING DATE: 1995-12-13
;  EARLIER APPLICATION NUMBER: 60/008,540
;  EARLIER FILING DATE: 1995-12-13
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
;   LENGTH: 933
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: DOMAIN
;   LOCATION: (659)..(918)
;   OTHER INFORMATION: minimal ligand binding domain
US-08-980-115-14

  Query Match             55.8%;  Score 765;  DB 3;  Length 933;
  Best Local Similarity   55.5%;  Pred. No. 5.6e-79;
  Matches  137;  Conservative   57;  Mismatches   53;  Indels    0;  Gaps    0;

Qy         12 PIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFRNLHV 71
              |  :|:| :||| |: ||||| :||: ::||:|||:|||||: ||||:|:||||||||:
Db        685 PPLINLLMSIEPDVIYAGHDNTKPDTSSSLLTSLNQLGERQLLSVVKWSKSLPGFRNLHI 744

Qy         72 DDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHL 131
              |||: :|||||| |||| :||||: :|: :|||||||: || || :|  || |: |  :
Db        745 DDQITLIQYSWMSLMVFGLGWRSYKHVSGQMLYFAPDLILNEQRMKESSFYSLCLTMWQI 804

Qy        132 SQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPT 191
               |||  ||:: :|||||| ||| : ||::||::|  |:|:| :||:|| : |  ::|
Db        805 PQEFVKLQVSQEEFLCMKVLLLLNTIPLEGLRSQTQFEEMRSSYIRELIKAIGLRQKGVV 864

Qy        192 SCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQVPKILSGK 251
              | |:||||||||||::  : ::|| :  :  |:|  :||:|||||:|:|: |:||||:|
Db        865 SSSQRFYQLTKLLDNLHDLVKQLHLYCLNTFIQSRALSVEFPEMMSEVIAAQLPKILAGM 924

Qy        252 VKPIYFH 258
              |||: ||
Db        925 VKPLLFH 931


RESULT 7
US-08-764-870-15
; Sequence 15, Application US/08764870
; Patent No. 6236946
;  GENERAL INFORMATION:
;    APPLICANT:  Scanlan, Thomas S
;    APPLICANT:  Baxter, John D
;    APPLICANT:  Fletterick, Robert J
;    APPLICANT:  Wagner, Richard L
;    APPLICANT:  Kushner, Peter J
;    APPLICANT:  Apriletti, James W
;    APPLICANT:  West, Brian
;    TITLE OF INVENTION:  Nuclear Receptor Ligands and Ligand
;    TITLE OF INVENTION:  Binding Domains
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooley Godward
;      STREET:  Five Palo Alto Square,  3000 El Camino Real
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/764,870
;      FILING DATE:  13-DEC-1996
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/008,540
;      FILING DATE:  13-DEC-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/008,543
;      FILING DATE:  13-DEC-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/008,606
;      FILING DATE:  14-DEC-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Nakamura, Jackie N
;      REGISTRATION NUMBER:  35,966
;      REFERENCE/DOCKET NUMBER:  UCAL-246/01US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650)843-5000
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  984 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-764-870-15

  Query Match             48.5%;  Score 665;  DB 3;  Length 984;
  Best Local Similarity   52.9%;  Pred. No. 2e-67;
  Matches  128;  Conservative   44;  Mismatches   70;  Indels    0;  Gaps    0;

Qy         17 VLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFRNLHVDDQMA 76
              ||| ||| :| ||:|:::||:   |||:|| |  :|:: |||||| ||||:|| ::||:
Db        741 VLENIEPEIVYAGYDSSKPDTAENLLSTLNRLAGKQMIQVVKWAKVLPGFKNLPLEDQIT 800

Qy         77 VIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHLSQEFG 136
              :|||||| |  ||: |||: : ||: |||||||||||| :||:| ||   |  | :| :|
Db        801 LIQYSWMCLSSFALSWRSYKHTNSQFLYFAPDLVFNEEKMHQSAMYELCQGMHQISLQFV 860

Qy        137 WLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPTSCSRR 196
               ||:| :|:  || ||| | || ||||:|  |:|:| ||||| :::      |     :|
Db        861 RLQLTFEEYTIMKVLLLLSTIPKDGLKSQAAFEEMRTNYIKELRKMVTKCPNNSGQSWQR 920

Qy        197 FYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQVPKILSGKVKPIY 256
              ||||||||||:  :  :| :| |    :|| : |:|| |: |||| |:||: ||  ||:|
Db        921 FYQLTKLLDSMHDLVSDLLEFCFYTFRESHALKVEFPAMLVEIISDQLPKVESGNAKPLY 980

Qy        257 FH 258
              ||
Db        981 FH 982


RESULT 8
US-08-980-115-15
; Sequence 15, Application US/08980115
; Patent No. 6266622
; GENERAL INFORMATION:
;  APPLICANT: Scanlan, Thomas S.
;  APPLICANT: Baxter, John D.
;  APPLICANT: Fletterick, Robert J.
;  APPLICANT: Wagner, Richard L.
;  APPLICANT: Kushner, Peter J.
;  APPLICANT: Apriletti, James W.
;  APPLICANT: West, Brian L.
;  APPLICANT: Shiau, Andrew K.
;  TITLE OF INVENTION: NUCLEAR RECEPTOR LIGANDS AND LIGAND BINDING DOMAINS
;  FILE REFERENCE: UCAL-246/02US
;  CURRENT APPLICATION NUMBER: US/08/980,115
;  CURRENT FILING DATE: 1997-11-26
```

```
;  EARLIER APPLICATION NUMBER: 08/764,870
;  EARLIER FILING DATE: 1996-12-13
;  EARLIER APPLICATION NUMBER: 60/008,606
;  EARLIER FILING DATE: 1995-12-14
;  EARLIER APPLICATION NUMBER: 60/008,543
;  EARLIER FILING DATE: 1995-12-13
;  EARLIER APPLICATION NUMBER: 60/008,540
;  EARLIER FILING DATE: 1995-12-13
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
;   LENGTH: 984
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: DOMAIN
;   LOCATION: (695)..(969)
;   OTHER INFORMATION: minimal ligand binding domain
US-08-980-115-15

  Query Match             48.5%;  Score 665;  DB 3;  Length 984;
  Best Local Similarity   52.9%;  Pred. No. 2e-67;
  Matches  128;  Conservative   44;  Mismatches   70;  Indels    0;  Gaps    0;

Qy         17 VLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFRNLHVDDQMA 76
Db        741 VLENIEPEIVYAGYDSSKPDTAENLLSTLNRLAGKQMIQVVKWAKVLPGFKNLPLEDQIT 800

Qy         77 VIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHLSQEFG 136
Db        801 LIQYSWMCLSSFALSWRSYKHTNSQFLYFAPDLVFNEEKMHQSAMYELCQGMHQISLQFV 860

Qy        137 WLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPTSCSRR 196
Db        861 RLQLTFEEYTIMKVLLLLSTIPKDGLKSQAAFEEMRTNYIKELRKMVTKCPNNSGQSWQR 920

Qy        197 FYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQVPKILSGKVKPIY 256
Db        921 FYQLTKLLDSMHDLVSDLLEFCFYTFRESHALKVEFPAMLVEIISDQLPKVESGNAKPLY 980

Qy        257 FH 258
Db        981 FH 982


RESULT 9
US-08-764-870-13
; Sequence 13, Application US/08764870
; Patent No. 6236946
;  GENERAL INFORMATION:
;    APPLICANT:  Scanlan, Thomas S
;    APPLICANT:  Baxter, John D
;    APPLICANT:  Fletterick, Robert J
;    APPLICANT:  Wagner, Richard L
;    APPLICANT:  Kushner, Peter J
;    APPLICANT:  Apriletti, James W
;    APPLICANT:  West, Brian
;    TITLE OF INVENTION:  Nuclear Receptor Ligands and Ligand
;    TITLE OF INVENTION:  Binding Domains
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooley Godward
;      STREET:  Five Palo Alto Square,  3000 El Camino Real
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/764,870
;      FILING DATE:  13-DEC-1996
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/008,540
;      FILING DATE:  13-DEC-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/008,543
;      FILING DATE:  13-DEC-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/008,606
;      FILING DATE:  14-DEC-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Nakamura, Jackie N
;      REGISTRATION NUMBER:  35,966
;      REFERENCE/DOCKET NUMBER:  UCAL-246/01US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650)843-5000
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  777 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-764-870-13

  Query Match             48.3%;  Score 662.5;  DB 3;  Length 777;
  Best Local Similarity   50.6%;  Pred. No. 2.7e-67;
  Matches  125;  Conservative   49;  Mismatches   72;  Indels    1;  Gaps    1;

Qy         12 PIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFRNLHV 71
Db        530 PTLVSLLEVIEPEVLYAGYDSSVPDSTWRIMTTLNMLGGRQVIAAVKWAKAIPGFRNLHL 589

Qy         72 DDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHL 131
Db        590 DDQMTLLQYSWMFLMAFALGWRSYRQSSANLLCFAPDLIINEQRMTLPCMYDQCKHMLYV 649

Qy        132 SQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPT 191
Db        650 SSELHRLQVSYEEYLCMKTLLLLSSVPKDGLKSQELFDEIRMTYIKELGKAIVKREGNSS 709

Qy        192 SCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQVPKILSGK 251
Db        710 QNWQRFYQLTKLLDSMHEVVENLLNYCFQTFLDKTM-SIEFPEMLAEIITNQIPKYSNGN 768

Qy        252 VKPIYFH 258
Db        769 IKKLLFH 775


RESULT 10
US-08-980-115-13
; Sequence 13, Application US/08980115
; Patent No. 6266622
; GENERAL INFORMATION:
;  APPLICANT: Scanlan, Thomas S.
;  APPLICANT: Baxter, John D.
;  APPLICANT: Fletterick, Robert J.
;  APPLICANT: Wagner, Richard L.
;  APPLICANT: Kushner, Peter J.
;  APPLICANT: Apriletti, James W.
;  APPLICANT: West, Brian L.
;  APPLICANT: Shiau, Andrew K.
;  TITLE OF INVENTION: NUCLEAR RECEPTOR LIGANDS AND LIGAND BINDING DOMAINS
;  FILE REFERENCE: UCAL-246/02US
;  CURRENT APPLICATION NUMBER: US/08/980,115
;  CURRENT FILING DATE: 1997-11-26
;  EARLIER APPLICATION NUMBER: 08/764,870
;  EARLIER FILING DATE: 1996-12-13
;  EARLIER APPLICATION NUMBER: 60/008,606
;  EARLIER FILING DATE: 1995-12-14
```

```
;  EARLIER APPLICATION NUMBER: 60/008,543
;  EARLIER FILING DATE: 1995-12-13
;  EARLIER APPLICATION NUMBER: 60/008,540
;  EARLIER FILING DATE: 1995-12-13
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
;   LENGTH: 777
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: DOMAIN
;   LOCATION: (506)..(762)
;   OTHER INFORMATION: minimal ligand binding domain
US-08-980-115-13

  Query Match             48.3%;  Score 662.5;  DB 3;  Length 777;
  Best Local Similarity   50.6%;  Pred. No. 2.7e-67;
  Matches  125;  Conservative   49;  Mismatches   72;  Indels    1;  Gaps    1;

Qy         12 PIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFRNLHV 71
              |  :::|| ||| |: ||:|:: |||   ::::|| || ||::  ||||||:|||||||:
Db        530 PTLVSLLEVIEPEVLYAGYDSSVPDSTWRIMTTLNMLGGRQVIAAVKWAKAIPGFRNLHL 589

Qy         72 DDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHL 131
              |||| ::||||| || ||:||||:   :: :| |||||: || ||     || ||  | ::
Db        590 DDQMTLLQYSWMFLMAFALGWRSYRQSSANLLCFAPDLIINEQRMTLPCMYDQCKHMLYV 649

Qy        132 SQEPGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPT 191
              | |   ||:: :|:|||| ||| | :| ||||:|: |||:|| ||||| : |  :  | :
Db        650 SSELHRLQVSYEEYLCMKTLLLLSSVPKDGLKSQELFDEIRMTYIKELGKAIVKREGNSS 709

Qy        192 SCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQVPKILSGK 251
                :||||||||||:  :   |  : |   :   | |::||||:||||: |:||  :|
Db        710 QNWQRFYQLTKLLDSMHEVVENLLNYCFQTFLDKTM-SIEFPEMLAEIITNQIPKYSNGN 768

Qy        252 VKPIYFH 258
              :| : ||
Db        769 IKKLLFH 775


RESULT 11
US-08-875-223-8
; Sequence 8, Application US/08875223
; Patent No. 6127175
;  GENERAL INFORMATION:
;    APPLICANT:  VIGNE, Emmanuelle
;    APPLICANT:  PERRICAUDET, Michel
;    APPLICANT:  DEDIEU, Jean-Francois
;    APPLICANT:  ORSINI, Cecile
;    APPLICANT:  YEH, Patrice
;    APPLICANT:  LATTA, Martine
;    APPLICANT:  PROST, Edouard
;    TITLE OF INVENTION:  CELLS FOR THE PRODUCTION OF RECOMBINANT
;    TITLE OF INVENTION:  ADENOVIRUSES
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Rhone-Poulenc Rorer Inc.
;      STREET:  500 Arcola Road, Mailstop 3C43
;      CITY:  Collegeville
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19426
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/875,223
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 95/00747
;      FILING DATE:  20-JAN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 95/06532
;      FILING DATE:  01-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 95/10541
;      FILING DATE:  08-SEP-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR WO FR96/00088
;      FILING DATE:  19-JAN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fehlner Esq., Paul F.
;      REGISTRATION NUMBER:  35,135
;      REFERENCE/DOCKET NUMBER:  ST95005G1-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (610) 454-3839
;      TELEFAX:  (610) 454-3808
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  534 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-875-223-8

  Query Match             47.0%;  Score 644.5;  DB 3;  Length 534;
  Best Local Similarity   51.3%;  Pred. No. 1.8e-65;
  Matches  122;  Conservative   46;  Mismatches   69;  Indels    1;  Gaps    1;

Qy         21 IEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFRNLHVDDQMAVIQY 80
              ||| |: ||:|:: |||   ::::|| || ||::  ||||||:|||||||:|||| ::||
Db          3 IEPEVLYAGYDSSVPDSTWRIMTTLNMLGGRQVIAAVKWAKAIPGFRNLHLDDQMTLLQY 62

Qy         81 SWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHLSQEFGWLQI 140
              ||| || ||:||||:   :: :| |||||: || ||     || ||  | ::| |    ||:
Db         63 SWMFLMAFALGWRSYRQSSANLLCFAPDLIINEQRMTLPCMYDQCKHMLYVSSELHRLQV 122

Qy        141 TPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPTSCSRRFYQL 200
              : :|:|||| ||| | :| ||||:|: |||:|| ||||| : |  :  | :   :|||||
Db        123 SYEEYLCMKTLLLLSSVPKDGLKSQELFDEIRMTYIKELGKAIVKREGNSSQNWQRFYQL 182

Qy        201 TKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQVPKILSGKVKPIYFH 258
              ||||||:  :   |  : |   :   | |::||||:||||: |:||  :| :| : ||
Db        183 TKLLDSMHEVVENLLNYCFQTFLDKTM-SIEFPEMLAEIITNQIPKYSNGNIKKLLFH 239


RESULT 12
US-09-091-042A-2
; Sequence 2, Application US/09091042A
; Patent No. 6455300
;  GENERAL INFORMATION:
;       APPLICANT: The Government of the United States of America
;                  as respresented by the Secretary
;                  Department of Health and Human Services
;                  Washington, D.C.
;                  Htun Ph.D., Han
;                  Hager Ph.D., Gordon L.
;       TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MONITORING
;                           DNA BINDING MOLECULES IN LIVING CELLS
;       NUMBER OF SEQUENCES: 9
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Needle & Rosenberg
;            STREET: 127 Peachtree Street, Suite 1200
;            CITY: Atlanta
;            STATE: Georgia
;            COUNTRY: USA
;            ZIP: 30303
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;          CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/091,042A
;            FILING DATE: 08-Jun-1998
;            CLASSIFICATION: <Unknown>
;          PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 60/008,373
;            FILING DATE: 08 Dec 1995
;          ATTORNEY/AGENT INFORMATION:
;            NAME: Selby, Elizabeth
;            REGISTRATION NUMBER: 38298
;            REFERENCE/DOCKET NUMBER: 14014.0183
;          TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 404-688-0770
;            TELEFAX: 404-688-9880
;       INFORMATION FOR SEQ ID NO: 2:
;          SEQUENCE CHARACTERISTICS:
;            LENGTH: 1070 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;          MOLECULE TYPE: protein
;          SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-091-042A-2

  Query Match           46.6%;  Score 639.5;  DB 4;  Length 1070;
  Best Local Similarity  49.4%;  Pred. No. 2e-64;
  Matches  122;  Conservative   48;  Mismatches   76;  Indels    1;  Gaps    1;

Qy         12 PIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFRNLHV 71
              |  :::|| ||| |: ||:|:: |||   ::::|| || ||::  |||||| | ||||:
Db        823 PTLVSLLEVIEPEVLYAGYDSSVPDSAWRIMTTLNMLGGRQVIAAVKWAKAILGLRNLHL 882

Qy         72 DDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHL 131
              |||| ::|||| || ||:|||:   :  :| |||||: || ||    || ||  |  :
Db        883 DDQMTLLQYSWMFLMAFALGWRSYRQSSGNLLCFAPDLIINEQRMSLPGMYDQCKHMLFV 942

Qy        132 SQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPT 191
              | |   ||:: :|:|||| ||| | :| :|||:|: |||:|| ||||| : |  :  | :
Db        943 SSELQRLQVSYEEYLCMKTLLLLSSVPKEGLKSQELFDEIRMTYIKELGKAIVKREGNSS 1002

Qy        192 SCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQVPKILSGK 251
                :||||||||||:  :   |  : |   :   | |::||||:||||: |:||  :|
Db       1003 QNWQRFYQLTKLLDSMHEVVENLLTYCFQTFLDKTM-SIEFPEMLAEIITNQIPKYSNGN 1061

Qy        252 VKPIYFH 258
              :| : ||
Db       1062 IKKLLFH 1068


RESULT 13
US-08-592-214A-24
; Sequence 24, Application US/08592214A
; Patent No. 5811536
;  GENERAL INFORMATION:
;    APPLICANT:  Yanofsky, Martin F.
;    TITLE OF INVENTION:  Cauliflower Floral Meristem Identify
;    TITLE OF INVENTION:  Genes and Methods of Using Same
;    NUMBER OF SEQUENCES:  33
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/592,214A
;      FILING DATE:  26-JAN-1996
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-UD 1927
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  284 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-592-214A-24

  Query Match           46.6%;  Score 638.5;  DB 2;  Length 284;
  Best Local Similarity  49.4%;  Pred. No. 3.5e-65;
  Matches  122;  Conservative   48;  Mismatches   76;  Indels    1;  Gaps    1;

Qy         12 PIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFRNLHV 71
              |  :::|| ||| |: ||:|:: |||   ::::|| || ||::  |||||| | ||||:
Db         37 PTLVSLLEVIEPEVLYAGYDSSVPDSAWRIMTTLNMLGGRQVIAAVKWAKAILGLRNLHL 96

Qy         72 DDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHL 131
              |||| ::|||| || ||:|||:   :  :| |||||: || ||    || ||  |  :
Db         97 DDQMTLLQYSWMFLMAFALGWRSYRQSSGNLLCFAPDLIINEQRMSLPCMYDQCKHMLFV 156

Qy        132 SQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPT 191
              | |   ||:: :|:|||| ||| | :| :|||:|: |||:|| ||||| : |  :  | :
Db        157 SSELQRLQVSYEEYLCMKTLLLLSSVPKEGLKSQELFDEIRMTYIKELGKAIVKREGNSS 216

Qy        192 SCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQVPKILSGK 251
                :||||||||||:  :   |  : |   :   | |::||||:||||: |:||  :|
Db        217 QNWQRFYQLTKLLDSMHEVVENLLTYCFQTFLDKTM-SIEFPEMLAEIITNQIPKYSNGN 275

Qy        252 VKPIYFH 258
              :| : ||
Db        276 IKKLLFH 282


RESULT 14
US-09-149-976-24
; Sequence 24, Application US/09149976
; Patent No. 6127123
;  GENERAL INFORMATION:
;    APPLICANT:  Yanofsky, Martin F.
;    TITLE OF INVENTION:  Cauliflower Floral Meristem Identity
;    TITLE OF INVENTION:  Genes and Methods of Using Same
;    NUMBER OF SEQUENCES:  33
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell & Flores LLP
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/149,976
;      FILING DATE:  09-SEP-1998
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/592,214
;      FILING DATE:  26-JAN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-UD 3291
```

```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  284 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-149-976-24

  Query Match             46.6%;  Score 638.5;  DB 3;  Length 284;
  Best Local Similarity   49.4%;  Pred. No. 3.5e-65;
  Matches  122;  Conservative   48;  Mismatches   76;  Indels    1;  Gaps    1;

Qy         12 PIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFRNLHV 71
              |  :::|| ||| |: ||:|:: |||   ::::|| || ||::  ||||||: | ||||:
Db         37 PTLVSLLEVIEPEVLYAGYDSSVPDSAWRIMTTLNMLGGRQVIAAVKWAKAILGLRNLHL 96

Qy         72 DDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHL 131
              |||| ::||||| || ||:||||:   :  :| |||||: || ||    || ||  |  :
Db         97 DDQMTLLQYSWMFLMAFALGWRSYRQSSGNLLCFAPDLIINEQRMSLPCMYDQCKHMLFV 156

Qy        132 SQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPT 191
              | |   ||:: :|:|||| ||| | :| :|||:|: |||:|| ||||| : |  :  | :
Db        157 SSELQRLQVSYEEYLCMKTLLLLSSVPKEGLKSQELFDEIRMTYIKELGKAIVKREGNSS 216

Qy        192 SCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQVPKILSGK 251
                :|||||||||||:  :   |  : |   :   | |::||||:||||: |:||  :|
Db        217 QNWQRFYQLTKLLDSMHEVVENLLTYCFQTFLDKTM-SIEFPEMLAEIITNQIPKYSNGN 275

Qy        252 VKPIYFH 258
              :| : ||
Db        276 IKKLLFH 282


RESULT 15
US-07-716-827C-5
; Sequence 5, Application US/07716827C
; Patent No. 5215916
;  GENERAL INFORMATION:
;    APPLICANT:  Simons Jr., Stoney S.
;    APPLICANT:  Yamamoto, K. R.
;    APPLICANT:  Chakraborti, P. K.
;    APPLICANT:  Garabedian, M. J.
;    TITLE OF INVENTION:  SUPER GLUCOCORTICOID RECEPTORS
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cushman, Darby & Cushman
;      STREET:  Eleventh floor, 1615 L Street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20036-5601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/716,827C
;      FILING DATE:  19910619
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Scott, Watson T.
;      REGISTRATION NUMBER:  26,581
;      REFERENCE/DOCKET NUMBER:  WTS/5683/84453
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)8613000
;      TELEFAX:  (202)822-8944
;      TELEX:  6714627 CUSH
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  795 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
US-07-716-827C-5

  Query Match             46.6%;  Score 638.5;  DB 1;  Length 795;
  Best Local Similarity   49.4%;  Pred. No. 1.6e-64;
  Matches  122;  Conservative   48;  Mismatches   76;  Indels    1;  Gaps    1;

Qy         12 PIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAKALPGFRNLHV 71
              |  :::|| ||| |: ||:|:: |||   ::::|| || ||::  ||||||: | ||||:
Db        548 PTLVSLLEVIEPEVLYAGYDSSVPDSAWRIMTTLNMLGGRQVIAAVKWAKAILGLRNLHL 607

Qy         72 DDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRMYSQCVRMRHL 131
              |||| ::||||| || ||:||||:   :  :| |||||: || ||    || ||  |  :
Db        608 DDQMTLLQYSWMFLMAFALGWRSYRQSSGNLLCFAPDLIINEQRMSLPCMYDQCKHMLFV 667

Qy        132 SQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDRIIACKRKNPT 191
              | |   ||:: :|:|||| ||| | :| :|||:|: |||:|| ||||| : |  :  | :
Db        668 SSELQRLQVSYEEYLCMKTLLLLSSVPKEGLKSQELFDEIRMTYIKELGKAIVKREGNSS 727

Qy        192 SCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIISVQVPKILSGK 251
                :|||||||||||:  :   |  : |   :   | |::||||:||||: |:||  :|
Db        728 QNWQRFYQLTKLLDSMHEVVENLLTYCFQTFLDKTM-SIEFPEMLAEIITNQIPKYSNGN 786

Qy        252 VKPIYFH 258
              :| : ||
Db        787 IKKLLFH 793


Search completed: October  9, 2003, 11:42:40
Job time : 20 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 9, 2003, 11:38:28 ; Search time 45 Seconds
(without alignments)
917.087 Million cell updates/sec

Title: US-09-687-609A-1
Perfect score: 1371
Sequence: 1 GSHMIEGYECQPIFLNVLEA.....SVQVPKILSGKVKPIYFHTQ 260

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1371	100.0	260	22	AAB97073	Rat androgen recep
2	1349.5	98.4	388	23	ABB83821	Human androgen rec
3	1349.5	98.4	452	20	AAY21627	Ligand binding dom
4	1349.5	98.4	839	23	ABG71292	Human androgen rec
5	1349.5	98.4	839	23	AAG68238	Fused androgen rec
6	1349.5	98.4	895	24	AAE32995	Macaca mulatta and
7	1349.5	98.4	895	24	AAE32996	Macaca mulatta and
8	1349.5	98.4	899	24	ABP72589	Mouse androgen rec
9	1349.5	98.4	902	10	AAP93110	Rat androgen recep
10	1349.5	98.4	902	10	AAP91006	Rat androgen recep
11	1349.5	98.4	907	24	ABG74229	Canine Androgen re
12	1349.5	98.4	918	20	AAY33491	Human androgen rec
13	1349.5	98.4	919	10	AAP90996	Human androgen rec
14	1349.5	98.4	919	18	AAW14783	Androgen receptor.
15	1349.5	98.4	919	21	AAY78914	Human androgen rec
16	1349.5	98.4	919	23	ABJ05582	Breast cancer-asso
17	1349.5	98.4	919	23	AAE19061	Human androgen rec
18	1349.5	98.4	919	24	ABJ19809	Androgen-independe
19	1346.5	98.2	918	12	AAR12223	Human androgen rec
20	1341.5	97.8	919	10	AAP93109	Human androgen rec
21	1338.5	97.6	902	12	AAR12224	Rat androgen recep
22	1326	96.7	252	24	ABU08036	Human steroid rece
23	1243	90.7	246	23	AAU75930	Androgen receptor
24	1179	86.0	250	22	AAG65959	Human androgen rec
25	937	68.3	630	12	AAR12230	TrpE/androgen rece
26	813.5	59.3	294	23	ABB83822	Human androgen rec
27	765	55.8	251	21	AAB26786	Human progesterone
28	765	55.8	251	23	ABJ15107	LBDG1 related prot
29	765	55.8	252	24	ABU08035	Human steroid rece
30	765	55.8	255	22	AAB97074	Rat progesterone r
31	765	55.8	314	24	ABP70513	Amino acid sequenc
32	765	55.8	933	20	AAY21621	Ligand binding dom
33	765	55.8	933	21	AAY97297	Human progesterone
34	728	53.1	240	23	AAE17863	Chain A of structu
35	706	51.5	630	23	ABB80765	Geneswitch regulat
36	706	51.5	630	23	ABB80766	Truncated GAL4 DBD
37	706	51.5	630	24	AAE35246	pGS1633 v.4.0 plas
38	706	51.5	630	24	AAE35248	Geneswitch regulat
39	706	51.5	654	23	ABB80764	Geneswitch regulat
40	706	51.5	654	24	AAE35245	pGLVp65 plasmid ge
41	702	51.2	648	24	AAE35247	Geneswitch regulat
42	666.5	48.6	483	24	ABP58319	Glucocorticoid rec
43	666	48.6	215	23	AAE17861	Chain A of structu
44	665.5	48.5	257	24	ABU08045	Mutant human gluco
45	665.5	48.5	777	24	ABU08024	Mutant human gluco

ALIGNMENTS

RESULT 1
AAB97073
ID AAB97073 standard; Protein; 260 AA.
XX
AC AAB97073;
XX
DT 31-JUL-2001 (first entry)
XX
DE Rat androgen receptor ligand binding domain.
XX
KW Rat; androgen receptor; AR; ligand binding domain; LBD; osteopathic;
KW crystallographic structure; AR-LBD; AR modulator; prostate cancer;
KW age related disease; osteoporosis; muscle wasting; libido; vasotropic;
KW protein coordinate data.
XX
OS Rattus sp.
XX
PN WO200127622-A1.
XX
PD 19-APR-2001.
XX
PF 13-OCT-2000; 2000WO-US28495.
XX
PR 14-OCT-1999; 99US-0159394.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Weinmann R, Einspahr HM, Krystek SR, Sack JS, Salvati ME;
PI Tokarski JS, Wang C, Attar RM;
XX
DR WPI; 2001-300222/31.
XX

PT New crystallographic structure of the Androgen receptor ligand binding
PT domain, useful for identifying modulators of androgen receptors -
XX
PS Example; Page 27; 83pp; English.
XX
CC The present sequence is the rat androgen receptor (AR) ligand-binding
CC domain (LBD). It is provided in an example illustrating an invention
CC relating to a new crystallographic structure of the AR-LBD. The
CC structure comprises either an AR-LBD and an AR-LBD ligand, or an
CC AR-LBD without an AR-LBD ligand (where the crystal diffracts to at
CC least 3 angstrom resolution and has a crystal stability within 5 % of
CC its unit cell dimensions). The AR-LBD is useful for identifying
CC modulators of the androgen receptor, which may be used for treating
CC prostate cancer and age related diseases such as osteoporosis, muscle
CC wasting and loss of libido.
XX
SQ Sequence 260 AA;

Query Match 100.0%; Score 1371; DB 22; Length 260;
Best Local Similarity 100.0%; Pred. No. 3.8e-135;
Matches 260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GSHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GSHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWA 60

Qy         61 KALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 KALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSR 120

Qy        121 MYSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 MYSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELD 180

Qy        181 RIIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEII 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 RIIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEII 240

Qy        241 SVQVPKILSGKVKPIYFHTQ 260
              ||||||||||||||||||||
Db        241 SVQVPKILSGKVKPIYFHTQ 260
```

RESULT 2
ABB83821
ID ABB83821 standard; Protein; 388 AA.
XX
AC ABB83821;
XX
DT 13-SEP-2002 (first entry)
XX
DE Human androgen receptor variant AR42 SEQ ID NO 2.
XX
KW Human; androgen receptor; receptor; AR; AR42; AR32; tumour; cancer;
KW steroid response element; prostate cancer; testis; gene therapy;
KW hormone therapy; cytostatic; contraceptive.
XX
OS Homo sapiens.
XX
PN EP1213300-A1.
XX
PD 12-JUN-2002.
XX
PF 26-OCT-2001; 2001EP-0250379.
XX
PR 30-NOV-2000; 2000DE-1061161.
XX
PA (SCHD) SCHERING AG.
XX
PI Ahrens-Fath I, Haendler B;
XX
DR WPI; 2002-530575/57.
DR N-PSDB; ABN85656.
XX
PT New nucleic acid encoding variant forms of androgen receptor; useful
PT for diagnosis and treatment of prostatic cancer, also related proteins
PT and antibodies -
XX
PS Claim 5; Page 9-11; 25pp; German.
XX
CC The invention relates to a nucleic acid (I) that encodes an androgen
CC receptor (AR) variant AR42 (ABN85656) or AR32 (ABN85657) or its
CC equivalents or a sequence that hybridises to them under stringent
CC conditions. Polypeptides (II, ABB83821-ABB83822) expressed by (I), can
CC bind androgens and other ligands and form heterodimers which can bind to
CC steroid response elements in target genes but do not induce activation,
CC so act as repressors of the known AR.(II) are used to raise specific
CC antibodies. (I) are used for recombinant production of (II) and as probes
CC for detecting (I)-related nucleic acid in tumour tissue. Also (I), (II),
CC a specific peptide and cells transfected with (I)-containing vectors, are
CC useful for identifying effectors of (II), specifically antiandrogenic
CC agents potentially useful for treating androgen-related diseases, e.g.
CC cancer of prostate or testis and as male contraceptives. (I) can be used
CC similarly in gene therapy. Antibodies to (II) are used to detect or
CC quantify (II) in tumour tissue, e.g. to determine if resistance to
CC hormone therapy is the result of altered expression of (II).
XX
SQ Sequence 388 AA;

Query Match 98.4%; Score 1349.5; DB 23; Length 388;
Best Local Similarity 99.6%; Pred. No. 1.2e-132;
Matches 258; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 189

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 249

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 309

Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 369

Qy        242 VQVPKILSGKVKPIYFHTQ 260
              ||||||||||||||||||||
Db        370 VQVPKILSGKVKPIYFHTQ 388
```

RESULT 3
AAY21627
ID AAY21627 standard; protein; 452 AA.
XX
AC AAY21627;
XX
DT 11-AUG-1999 (first entry)
XX
DE Ligand binding domain of nuclear receptor hAR.
XX
KW Thyroid hormone receptor; aromatic compound; ligand binding domain;
KW alpha-glcerophosphate dehydrogenase; cardiac; obesity; triglyceride;
KW plasma cholesterol; anti-hypertriglyceridaemic; atherosclerosis; GPDH;
KW thyroid hormone replacement therapy; nuclear receptor.
XX
OS Homo sapiens.
XX
PN WO9926966-A2.
XX
PD 03-JUN-1999.
XX
PF 25-NOV-1998; 98WO-US25296.

XX
PR 26-NOV-1997; 97US-0980115.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Apriletti JW, Baxter JD, Fletterick RJ, Kushner PJ;
PI Scanlan TS, Shiau AK, Wagner RL, West BL;
XX
DR WPI; 1999-357810/30.
XX
PT Modulating activity of a thyroid hormone receptor
XX
PS Disclosure; Fig 3H-R; 447pp; English.
XX
CC The invention relates to a method for modulating activity of a thyroid
CC hormone receptor that comprises administration of an aromatic compound
CC which fits spatially and preferentially into a thyroid hormone ligand
CC binding domain. The aromatic compound (of a specified formula) can be
CC used to increase alpha-glcerophosphate dehydrogenase (GPDH) levels, at
CC levels which do not significantly modify cardiac GPDH levels and are
CC indicated in the treatment of obesity. The compound also lower total
CC plasma cholesterol and triglyceride levels and can be used as anti-
CC hypertriglyceridaemic agents. The compound may also be used for treating
CC atherosclerosis and may be indicated in thyroid hormone replacement
CC therapy in patients with compromised cardiac function. Sequences
CC AAY21621- 636 amino acid sequences of ligand binding domains of several
CC members of the nuclear receptor superfamily.
XX
SQ Sequence 452 AA;

Query Match 98.4%; Score 1349.5; DB 20; Length 452;
Best Local Similarity 99.6%; Pred. No. 1.5e-132;
Matches 258; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 253

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 313

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 373

Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 433

Qy        242 VQVPKILSGKVKPIYFHTQ 260
              |||||||||||||||||||
Db        434 VQVPKILSGKVKPIYFHTQ 452
```

RESULT 4
ABG71292
ID ABG71292 standard; Protein; 839 AA.
XX
AC ABG71292;
XX
DT 08-JAN-2003 (first entry)
XX
DE Human androgen receptor.
XX
KW Pharmacological; receptor; endocrine disrupting chemical; EDC;
KW simultaneous determination; ligand binding activity; human;
KW androgen receptor.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Bacteria.
OS Synthetic.
XX
PN JP2002243739-A.
XX
PD 28-AUG-2002.
XX
PF 21-FEB-2001; 2001JP-0045420.
XX
PR 21-FEB-2001; 2001JP-0045420.
XX
PA (TOYM) TOYOBO KK.
XX
DR WPI; 2002-744851/81.
DR N-PSDB; ABS56325.
XX
PT A method and a kit for simultaneous determination of binding activity
PT of ligand with 2 or more receptors using their corresponding tracers -
XX
PS Example 3; Page 13-15; 22pp; Japanese.
XX
CC The present invention relates to a new method for simultaneous
CC determination of binding activity of a ligand with 2 or more receptors
CC using their corresponding tracers, and a kit for the determination.
CC The invention can be used for screening of ligands with pharmacological
CC activity via receptor and endocrine disrupting chemicals (EDCs). The
CC invention is advantageous as it allows simultaneous determination of
CC binding activity of ligand. The present amino acid sequence represents
CC the human androgen receptor as described in the invention.
XX
SQ Sequence 839 AA;

Query Match 98.4%; Score 1349.5; DB 23; Length 839;
Best Local Similarity 99.6%; Pred. No. 3.6e-132;
Matches 258; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        582 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 640

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        641 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 700

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        701 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 760

Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        761 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 820

Qy        242 VQVPKILSGKVKPIYFHTQ 260
              |||||||||||||||||||
Db        821 VQVPKILSGKVKPIYFHTQ 839
```

RESULT 5
AAG68238
ID AAG68238 standard; Protein; 839 AA.
XX
AC AAG68238;
XX
DT 08-FEB-2002 (first entry)
XX
DE Fused androgen receptor (AR) protein SEQ ID NO:11.
XX
KW Human; androgen receptor; AR; fused androgen receptor protein;
KW fusion androgen receptor protein; sugar-combining protein;
KW maltose-combining protein.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Unidentified.
XX
PN JP2001252080-A.
XX

PD 18-SEP-2001.
XX
PF 13-MAR-2000; 2000JP-0069030.
XX
PR 13-MAR-2000; 2000JP-0069030.
XX
PA (TOYM) TOYOBO KK.
XX
DR WPI; 2002-029658/04.
DR N-PSDB; ABA01683.
XX
PT New polypeptide for screening drugs, comprises an androgen receptor
PT protein fused with a sugar-combining protein -
XX
PS Claim 8; Page 11-13; 16pp; Japanese.
XX
CC The present invention describes a fused androgen receptor (AR) protein
CC prepared by fusing an androgen receptor protein with a sugar-combining
CC protein. Also described are: (1) a gene encoding the above fused AR
CC protein; (2) a vector carrying the above gene; (3) a transformant in
CC which the above vector is introduced to a microbe and a gene encoding
CC the fused AR protein is expressed; (4) the preparation of a fused AR
CC protein in which the above transformant is cultured and the fused AR
CC protein is collected from the resultant culture; and (5) a reagent for
CC detecting the presence of interaction with a ligand containing the
CC above fused AR protein, a solvent for dissolving chemical substances
CC and a dilutent liquid of the dissolved chemical substances. The fusion
CC protein can be used for screening drugs. The present sequence represents
CC the AR protein and sugar-combining protein fusion protein from the
CC present invention.
XX
SQ Sequence 839 AA;

Query Match 98.4%; Score 1349.5; DB 23; Length 839;
Best Local Similarity 99.6%; Pred. No. 3.6e-132;
Matches 258; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        582 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 640

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        641 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 700

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        701 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 760

Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        761 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 820

Qy        242 VQVPKILSGKVKPIYFHTQ 260
              ||||||||||||||||||
Db        821 VQVPKILSGKVKPIYFHTQ 839
```

RESULT 6
AAE32995
ID AAE32995 standard; Protein; 895 AA.
XX
AC AAE32995;
XX
DT 02-APR-2003 (first entry)
XX
DE Macaca mulatta androgen receptor (rhAR).
XX
KW Androgen receptor; rhAR protein; bone formation; atherosclerosis; acne;
KW myoanabolism; sarcopaenia; benign prostatic hyperplasia; hypogonadism;
KW lipid metabolism; hirsutism; prostate disease; hippocampal function;
KW cancer; rhesus monkey; single nucleotide polymorphism; SNP; therapy.
XX
OS Macaca mulatta.
XX
FH Key Location/Qualifiers
FT Misc-difference 210
FT /note= "This residue changes to Gly due to single
FT nucleotide polymorphism (SNP)"
FT Domain 535..600
FT /note= "DNA binding domain"
XX
PN WO200290529-A1.
XX
PD 14-NOV-2002.
XX
PF 03-MAY-2002; 2002WO-US14175.
XX
PR 08-MAY-2001; 2001US-289573P.
XX
PA (MERI) MERCK & CO INC.
XX
PI Towler DA, Chen F;
XX
DR WPI; 2003-103516/09.
DR N-PSDB; AAD50740.
XX
PT New DNA encoding Macaca mulatta androgen receptor (rhAR) protein,
PT useful for screening rhAR agonists and/or antagonists, and in
PT identifying tissue selective androgen compounds such as those active in
PT bone formation or myoanabolism -
XX
PS Claim 25; Page 38; 84pp; English.
XX
CC The invention relates to a DNA encoding Macaca mulatta androgen receptor
CC (rhAR) protein. Nucleic acid molecules of the invention are useful for
CC screening agonists and/or antagonists of rhAR and in identifying tissue
CC selective androgen compounds including those active in bone formation,
CC myoanabolism, treatment of sarcopaenia, benign prostatic hyperplasia,
CC acne, hirsutism, male hypogonadism, atherosclerosis, relief of post-
CC menopausal symptoms, treatment and prevention of prostate or breast
CC cancer and management of lipids. AR modulators may be used to regulate
CC development, production and maintenance of bone and muscle, in the
CC treatment of prostate disease and in regulation of lipid metabolism
CC and hippocampal function. The present sequence is Macaca mulatta
CC (rhesus monkey) androgen receptor.
XX
SQ Sequence 895 AA;

Query Match 98.4%; Score 1349.5; DB 24; Length 895;
Best Local Similarity 99.6%; Pred. No. 3.9e-132;
Matches 258; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        638 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 696

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        697 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 756

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        757 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 816

Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        817 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 876

Qy        242 VQVPKILSGKVKPIYFHTQ 260
              ||||||||||||||||||
Db        877 VQVPKILSGKVKPIYFHTQ 895
```

RESULT 7
AAE32996

```
ID   AAE32996 standard; Protein; 895 AA.
XX
AC   AAE32996;
XX
DT   02-APR-2003  (first entry)
XX
DE   Macaca mulatta androgen receptor (rhAR) allelic variant protein.
XX
KW   Androgen receptor; rhAR protein; bone formation; atherosclerosis; acne;
KW   myoanabolism; sarcopaenia; benign prostatic hyperplasia; hypogonadism;
KW   lipid metabolism; hirsutism; prostate disease; hippocampal function;
KW   cancer; rhesus monkey; single nucleotide polymorphism; SNP; therapy.
XX
OS   Macaca mulatta.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 210
FT                   /note= "wild-type Glu is changed to Gly due to single
FT                   nucleotide polymorphism (SNP)"
FT   Domain          535..600
FT                   /note= "DNA binding domain"
XX
PN   WO200290529-A1.
XX
PD   14-NOV-2002.
XX
PF   03-MAY-2002; 2002WO-US14175.
XX
PR   08-MAY-2001; 2001US-289573P.
XX
PA   (MERI ) MERCK & CO INC.
XX
PI   Towler DA,  Chen F;
XX
DR   WPI; 2003-103516/09.
DR   N-PSDB; AAD50741.
XX
PT   New DNA encoding Macaca mulatta androgen receptor (rhAR) protein,
PT   useful for screening rhAR agonists and/or antagonists, and in
PT   identifying tissue selective androgen compounds such as those active in
PT   bone formation or myoanabolism  -
XX
PS   Claim 55; Page 46; 84pp; English.
XX
CC   The invention relates to a DNA encoding Macaca mulatta androgen receptor
CC   (rhAR) protein. Nucleic acid molecules of the invention are useful for
CC   screening agonists and/or antagonists of rhAR and in identifying tissue
CC   selective androgen compounds including those active in bone formation,
CC   myoanabolism, treatment of sarcopaenia, benign prostatic hyperplasia,
CC   acne, hirsutism, male hypogonadism, atherosclerosis, relief of post-
CC   menopausal symptoms, treatment and prevention of prostate or breast
CC   cancer and management of lipids. AR modulators may be used to regulate
CC   development, production and maintenance of bone and muscle, in the
CC   treatment of prostate disease and in regulation of lipid metabolism
CC   and hippocampal function. The present sequence is Macaca mulatta
CC   (rhesus monkey) androgen receptor allelic variant protein.
XX
SQ   Sequence   895 AA;

  Query Match             98.4%;  Score 1349.5;  DB 24;  Length 895;
  Best Local Similarity   99.6%;  Pred. No. 3.9e-132;
  Matches  258;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        638 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 696

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        697 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 756

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        757 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 816

Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        817 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 876

Qy        242 VQVPKILSGKVKPIYFHTQ 260
              ||||||||||||||||||
Db        877 VQVPKILSGKVKPIYFHTQ 895


RESULT 8
ABP72589
ID   ABP72589 standard; Protein; 899 AA.
XX
AC   ABP72589;
XX
DT   29-MAY-2003  (first entry)
XX
DE   Mouse androgen receptor.
XX
KW   Androgen receptor; receptor; ARKO; mouse; transgenic mouse;
KW   knockout animal; cancer; tumour.
XX
OS   Mus musculus.
XX
PN   WO2003012394-A2.
XX
PD   13-FEB-2003.
XX
PF   29-JUL-2002; 2002WO-US24234.
XX
PR   27-JUL-2001; 2001US-308356P.
PR   17-MAY-2002; 2002US-381309P.
XX
PA   (UYRP ) UNIV ROCHESTER.
XX
PI   Chang C,  Yeh S;
XX
DR   WPI; 2003-247910/25.
DR   N-PSDB; ABZ81677, ABZ81679.
XX
PT   Composition for generating androgen receptor knockout mice for studying
PT   the role of an androgen receptor in cancer, comprises a cell having a
PT   disrupted androgen receptor gene -
XX
PS   Disclosure; Page 75; 99pp; English.
XX
CC   The present sequence is that of the murine androgen receptor (AR).
CC   The invention relates to compositions and methods for disrupting an
CC   AR gene and for studying the role of the AR in cancer.  Compositions
CC   comprising a cell having a disrupted AR gene are claimed.  The cell
CC   is especially an embryonic stem cell, an embryonic germ cell, a
CC   breast cell, a breast cancer cell, an ovary cell, an ovary cancer
CC   cell, a prostate cell, a testis cell, a bone cell, a brain cell, a
CC   neural cell or a muscle cell.  Also claimed is a transgenic mammal,
CC   especially a mouse, comprising a disrupted AR gene, especially a
CC   disrupted AR gene lacking exon 2 of the AR gene, where the
CC   disruption is produced by the action of a cre recombinase under the
CC   control of an inducible promoter specific for breast, ovary, neural,
CC   bone, testis, liver or prostate.  Also claimed is a method of
CC   determining the effect of a steroid on the AR by incubating the
CC   steroid with an AR disrupted cell line and assaying its effect.  A
CC   claimed method of evaluating treatment for cancer xenografts
CC   involves injecting a cell that has a disrupted AR locus into an
CC   ovarectomised mouse, especially a nude mouse, where the xenograft
CC   comprises a cell from a breast cancer or ovarian cancer tissue line.
CC   Tumour formation can be evaluated in an androgen receptor knockout
CC   (ARKO) mouse by injecting the cancer-causing agent into the mouse.
XX
SQ   Sequence   899 AA;

  Query Match             98.4%;  Score 1349.5;  DB 24;  Length 899;
```

```
  Best Local Similarity   99.6%;  Pred. No. 4e-132;
  Matches  258;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        642 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 700

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        701 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 760

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        761 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 820

Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        821 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 880

Qy        242 VQVPKILSGKVKPIYFHTQ 260
              ||||||||||||||||||
Db        881 VQVPKILSGKVKPIYFHTQ 899


RESULT 9
AAP93110
ID   AAP93110 standard; protein; 902 AA.
XX
AC   AAP93110;
XX
DT   25-MAR-2003  (updated)
DT   19-MAR-1990  (first entry)
XX
DE   Rat androgen receptor.
XX
KW   Rat androgen receptor; monoclonal antibody; ploYclonal antibody; cancer.
XX
OS   Rattus rattus.
XX
PN   WO8909791-A.
XX
PD   19-OCT-1989.
XX
PF   13-APR-1989;   89WO-US01548.
XX
PR   14-APR-1988;   88US-0182646.
XX
PA   (UYNC-) UNIV OF N CAROLINA.
XX
PI   French PS,  Wilson EM,  Joseph DR,  Lubahn DB;
XX
DR   WPI; 1989-324206/44.
DR   N-PSDB; AAN91773.
XX
PT   DNA encoding androgen receptor protein - useful for transforming
PT   eukaryotic hosts for protein expression and subsequent antibody prodn.
XX
PS   Disclosure; Fig. 5; 41pp; English.
XX
CC   Androgen receptor protein (AR) is used to produce mono- or poly-clonal
CC   antibodies.  These are used for the detection and quantification of AR in
CC   the presence of endogenous androgen, as androgen will not interfere with
CC   binding.  They may be used in assays to determine and quantify cellular
CC   distribution of AR in tumour tissue, and are esp. useful for evaluating
CC   prostate cancers to determine responsiveness to androgen withdrawal
CC   therapy.
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence   902 AA;

  Query Match             98.4%;  Score 1349.5;  DB 10;  Length 902;
  Best Local Similarity   99.6%;  Pred. No. 4e-132;
  Matches  258;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        645 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 703

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        704 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 763

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        764 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 823

Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        824 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 883

Qy        242 VQVPKILSGKVKPIYFHTQ 260
              ||||||||||||||||||
Db        884 VQVPKILSGKVKPIYFHTQ 902


RESULT 10
AAP91006
ID   AAP91006 standard; protein; 902 AA.
XX
AC   AAP91006;
XX
DT   25-MAR-2003  (updated)
DT   28-FEB-1990  (first entry)
XX
DE   Rat androgen receptor DNA clone.
XX
KW   Androgen receptor; TR2 polypeptide;
XX
OS   Rat.
XX
FH   Key             Location/Qualifiers
FT   Region          1..902
FT                   /*tag=  a
FT                   /product=98 kD polypeptide
FT   Region          170..902
FT                   /*tag=  b
FT                   /product=79 kD polypeptide
XX
PN   WO8909223-A.
XX
PD   05-OCT-1989.
XX
PF   24-MAR-1989;   89WO-JP01238.
XX
PR   30-MAR-1988;   88US-0176107.
PR   05-OCT-1988;   88US-0253807.
PR   21-FEB-1989;   89US-0312763.
XX
PA   (ARCH-) ARCH DEV CORP.
XX
PI   Liao S,  Chang C;
XX
DR   WPI; 1989-309501/42.
DR   N-PSDB; AAN91578.
XX
PT   New DNA encoding new androgen receptor and TR2 polypeptide(s) - able
PT   to bind DNA, and derived antibodies, useful for receptor assay and
PT   purification.
XX
PS   Claim 8; Fig 3; 60pp; English.
XX
CC   The protein is used to raise antibodies for receptor assays and for
CC   affinity purification.
CC   The 98 kD product starts at the first Met codon; the 79 kD product
CC   starts from the second.
CC   (Updated on 25-MAR-2003 to correct PF field.)
```

```
CC   (Updated on 25-MAR-2003 to correct PR field.)
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence   902 AA;

  Query Match             98.4%;  Score 1349.5;  DB 10;  Length 902;
  Best Local Similarity   99.6%;  Pred. No. 4e-132;
  Matches  258;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        645 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 703

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        704 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 763

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        764 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 823

Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        824 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 883

Qy        242 VQVPKILSGKVKPIYFHTQ 260
              ||||||||||||||||||
Db        884 VQVPKILSGKVKPIYFHTQ 902


RESULT 11
ABG74229
ID   ABG74229 standard; Protein; 907 AA.
XX
AC   ABG74229;
XX
DT   16-APR-2003  (first entry)
XX
DE   Canine Androgen receptor.
XX
KW   Dog; receptor; androgen receptor; AR; cytostatic; osteopathic;
KW   neuroleptic; canine disorder; cancer; mood disorder; temper disorder;
KW   frailty; muscular degeneration; bone loss.
XX
OS   Canis familiaris.
XX
PN   US2002161194-A1.
XX
PD   31-OCT-2002.
XX
PF   09-NOV-2001; 2001US-0008739.
XX
PR   09-NOV-2000; 2000US-247373P.
XX
PA   (CAST/) CASTLEBERRY T A.
PA   (LUBB/) LU B.
PA   (OWEN/) OWEN T A.
PA   (SMOC/) SMOCK S L.
XX
PI   Castleberry TA,  Lu B,  Owen TA,  Smock SL;
XX
DR   WPI; 2003-209230/20.
DR   N-PSDB; ABX16536.
XX
PT   New isolated canine androgen receptor protein, useful for diagnosing,
PT   classifying, prognosing and/or treating canine diseases (e.g. cancer)
PT   and in studying the molecular and biochemical actions of androgens in
PT   canine bone  -
XX
PS   Claim 1; Fig 1; 20pp; English.
XX
CC   The invention relates to an isolated proteinaceous molecule having canine
CC   androgen receptor (AR) activity (appearing as ABG74229), optionally
CC   having one or more conservative substitutions. Also included are
CC   an isolated DNA molecule encoding the above proteinaceous molecule,
CC   a recombinant expression vector comprising the above DNA molecule,
CC   a cell transformed by the expression vector cited above, producing the
CC   proteinaceous molecule by culturing the cell cited above, a specific
CC   binding partner that selectively binds to the proteinaceous molecule,
CC   discovering ligands for the canine AR protein, and discovering modulators
CC   of canine AR protein activity. The proteinaceous molecule is useful in
CC   veterinary purposes for diagnosing, classifying, prognosing and/or
CC   treating canine disorders (e.g. cancer, mood or temper disorders, and
CC   frailty associated with muscular degeneration or bone loss) which maybe
CC   characterised as related to the interaction between a cell receptor and
CC   its specific ligand, and in further studying the molecular and
CC   biochemical actions of androgens in canine bone. The DNA molecule
CC   encoding the proteinaceous molecule is used in DNA hybridisation
CC   processes to locate the canine gene position and/or the position of any
CC   related gene family in a chromosomal map. These may also be used for
CC   identifying canine gene disorders at the DNA level and used as gene
CC   markers for identifying neighbouring genes and their disorders.
CC   The present sequence represents the canine androgen receptor.
XX
SQ   Sequence   907 AA;

  Query Match             98.4%;  Score 1349.5;  DB 24;  Length 907;
  Best Local Similarity   99.6%;  Pred. No. 4e-132;
  Matches  258;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        650 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 708

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        709 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 768

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        769 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 828

Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        829 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 888

Qy        242 VQVPKILSGKVKPIYFHTQ 260
              ||||||||||||||||||
Db        889 VQVPKILSGKVKPIYFHTQ 907


RESULT 12
AAY33491
ID   AAY33491 standard; Protein; 918 AA.
XX
AC   AAY33491;
XX
DT   19-JAN-2000  (first entry)
XX
DE   Human androgen receptor protein.
XX
KW   Proapoptotic; dependence domain; p75NTR; androgen receptor; DCC;
KW   huntingtin polypeptide; Machado-Joseph disease; SCA1; SCA2; SCA6;
KW   atrophin-1; cell death; apoptosis; Huntington's disease; head trauma;
KW   Alzheimer's disease; Kennedy's disease; spinocerebellar ataxia; stroke;
KW   dentatorubropallidoluysian atrophy; cell proliferation; cell survival;
KW   neoplastic; malignant; autoimmune; fibrotic.
XX
OS   Homo sapiens.
XX
PN   WO9945944-A1.
XX
PD   16-SEP-1999.
XX
PF   11-MAR-1999;   99WO-US05250.
XX
```

```
PR   12-MAR-1998;   98US-0041886.
XX
PA   (BURN-) BURNHAM INST.
XX
PI   Bredesen DE,  Rabizadeh S;
XX
DR   WPI; 1999-561617/47.
DR   N-PSDB; AAZ23424.
XX
PT   New proapoptotic dependence peptides, used to develop products for
PT   treating, e.g. Alzheimer's disease -
XX
PS   Disclosure; Page 90-93; 199pp; English.
XX
CC   This invention describes novel pure proapoptotic dependence peptides
CC   which comprise a sequence of an active dependence domain selected from
CC   dependence polypeptides consisting of p75NTR, androgen receptor, DCC,
CC   huntingtin polypeptide, Machado-Joseph disease gene product, SCA1, SCA2,
CC   SCA6 and atrophin-1 polypeptide. The proapoptotic peptides are capable
CC   of inducing cell death and can be used to develop products to mediate or
CC   inhibit apoptosis. The methods can be used for reducing the severity of
CC   a proapoptotic dependence domain mediated pathological conditions e.g.
CC   Huntington's disease, Alzheimer's disease, Kennedy's disease,
CC   Spinocerebellar ataxias, dentatorubropallidoluysian atrophy,
CC   Machado-Joseph disease, stroke or head trauma. They can also be used for
CC   reducing the severity of a pathological condition mediated by upregulated
CC   cell proliferation or cell survival e.g. neoplastic, malignant,
CC   autoimmune or fibrotic conditions. This sequence represents a human
CC   androgen receptor described in the method of the invention.
XX
SQ   Sequence   918 AA;

  Query Match             98.4%;  Score 1349.5;  DB 20;  Length 918;
  Best Local Similarity   99.6%;  Pred. No. 4.1e-132;
  Matches  258;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 719

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 779

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 839

Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 899

Qy        242 VQVPKILSGKVKPIYFHTQ 260
              ||||||||||||||||||
Db        900 VQVPKILSGKVKPIYFHTQ 918


RESULT 13
AAP90996
ID   AAP90996 standard; protein; 919 AA.
XX
AC   AAP90996;
XX
DT   25-MAR-2003  (updated)
DT   28-FEB-1990  (first entry)
XX
DE   Human androgen receptor DNA clone.
XX
KW   Androgen receptor; TR2 polypeptide;
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Region          1..919
FT                   /*tag=  a
FT                   /product=98 kD polypeptide
FT   Region          185..919
FT                   /*tag=  b
FT                   /product=79 kD polypeptide
XX
PN   WO8909223-A.
XX
PD   05-OCT-1989.
XX
PF   24-MAR-1989;   89WO-JP01238.
XX
PR   30-MAR-1988;   88US-0176107.
PR   05-OCT-1988;   88US-0253807.
PR   21-FEB-1989;   89US-0312763.
XX
PA   (ARCH-) ARCH DEV CORP.
XX
PI   Liao S,  Chang C;
XX
DR   WPI; 1989-309501/42.
DR   N-PSDB; AAN91577.
XX
PT   New DNA encoding new androgen receptor and TR2 polypeptide(s) - able
PT   to bind DNA, and derived antibodies, useful for receptor assay and
PT   purification.
XX
PS   Claim 8; Fig 3; 60pp; English.
XX
CC   The protein is used to raise antibodies for receptor assays and for
CC   affinity purification.
CC   The 98 kD product starts at the first Met codon; the 79 kD product
CC   starts from the second.
CC   (Updated on 25-MAR-2003 to correct PF field.)
CC   (Updated on 25-MAR-2003 to correct PR field.)
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence   919 AA;

  Query Match             98.4%;  Score 1349.5;  DB 10;  Length 919;
  Best Local Similarity   99.6%;  Pred. No. 4.1e-132;
  Matches  258;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 720

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 780

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 840

Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 900

Qy        242 VQVPKILSGKVKPIYFHTQ 260
              ||||||||||||||||||
Db        901 VQVPKILSGKVKPIYFHTQ 919


RESULT 14
AAW14783
ID   AAW14783 standard; Protein; 919 AA.
XX
AC   AAW14783;
XX
DT   22-JUN-1997  (first entry)
XX
```

```
DE   Androgen receptor.
XX
KW   Androgen receptor; acidic fibroblast growth factor; aFGF;
KW   antisense; benign prostatic hyperplasia; prostate cancer; therapy.
XX
OS   Homo sapiens.
XX
PN   WO9711170-A1.
XX
PD   27-MAR-1997.
XX
PF   20-SEP-1996;   96WO-US15081.
XX
PR   20-SEP-1995;   95US-0004018.
XX
PA   (WORC-) WORCESTER FOUND BIOMEDICAL RES.
XX
PI   Zamecnik PA;
XX
DR   WPI; 1997-202879/18.
DR   N-PSDB; AAT63407.
XX
PT   Oligonucleotide(s) antisense to human androgen receptor and acidic
PT   FGF genes - used to inhibit gene expression, for the treatment of
PT   benign prostatic hyperplasia
XX
PS   Disclosure; Page 22-28; 51pp; English.
XX
CC   Human androgen receptor (AAW14783) binds testosterone and, acting
CC   at the transcriptional level, regulates the growth of normal
CC   prostatic cells.  Antisense oligonucleotides (see also AAT63200,
CC   AAT63404-05) based on an androgen receptor cDNA clone (see also
CC   AAT63407) can be used to prevent androgen receptor gene expression,
CC   thereby inhibiting the growth or survival of prostatic cells for
CC   the treatment of benign prostatic hyperplasia and prostate cancer.
XX
SQ   Sequence   919 AA;

  Query Match             98.4%;  Score 1349.5;  DB 18;  Length 919;
  Best Local Similarity   99.6%;  Pred. No. 4.1e-132;
  Matches  258;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 720

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 780

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 840

Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 900

Qy        242 VQVPKILSGKVKPIYFHTQ 260
              ||||||||||||||||||
Db        901 VQVPKILSGKVKPIYFHTQ 919


RESULT 15
AAY78914
ID   AAY78914 standard; protein; 919 AA.
XX
AC   AAY78914;
XX
DT   23-MAY-2000  (first entry)
XX
DE   Human androgen receptor (AR) amino acid sequence.
XX
KW   Androgen receptor; AR; androgen-independent activation; inhibitor;
KW   cancer; benign prostatic hyperplasia; hirsutism; androgenic alopecia;
KW   acne; breast cancer; Kennedy disease; prostate cancer.
XX
OS   Homo sapiens.
XX
PN   WO200001813-A2.
XX
PD   13-JAN-2000.
XX
PF   30-JUN-1999;   99WO-CA00604.
XX
PR   30-JUN-1998;   98US-0091871.
XX
PA   (UYBR-) UNIV BRITISH COLUMBIA.
XX
PI   Sadar MD,  Bruchovsky N,  Gout PW,  Snoek R,  Mawji NR;
XX
DR   WPI; 2000-182113/16.
XX
PT   Novel non-androgen ligand binding peptides for inhibiting
PT   androgen-independent activation of androgen receptor, used for
PT   screening compounds and for treatment of androgen-mediated diseases
PT   such as prostate cancer  -
XX
PS   Disclosure; Page 7; 32pp; English.
XX
CC   This sequence represents the human androgen receptor (AR) amino acid
CC   sequence. The invention relates to a fragment of the AR corresponding to
CC   amino acids 234-391 (see AAY78913). The fragment mediates
CC   androgen-independent activation of the AR. The androgen receptor acts as
CC   a transcription factor, regulating the expression of certain
CC   androgen-responsive genes. Interaction of the AR with the protein kinase
CC   A signal transduction pathway involves interaction with the androgen
CC   independent region. The AR fragment and peptides derived from it can be
CC   used as agents for inhibiting androgen independent activation of the
CC   androgen receptor, as activation domains, and as a tool for screening
CC   for compounds which affect androgen-independent activation of the AR.
CC   The peptides, when used in combination with androgen deprivation,
CC   effectively limit androgen mediated disease progression. These diseases
CC   include cancer, benign prostatic hyperplasia, hirsutism, androgenic
CC   alopecia, acne, breast cancer, Kennedy disease, and especially prostate
CC   cancer. The peptides and nucleic acids encoding them, are especially used
CC   for the treatment of androgen-mediated diseases, especially prostate
CC   tumours in patients deprived of androgen.
XX
SQ   Sequence   919 AA;

  Query Match             98.4%;  Score 1349.5;  DB 21;  Length 919;
  Best Local Similarity   99.6%;  Pred. No. 4.1e-132;
  Matches  258;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          2 SHMIEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 61
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 SH-IEGYECQPIFLNVLEAIEPGVVCAGHDNNQPDSFAALLSSLNELGERQLVHVVKWAK 720

Qy         62 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 121
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ALPGFRNLHVDDQMAVIQYSWMGLMVFAMGWRSFTNVNSRMLYFAPDLVFNEYRMHKSRM 780

Qy        122 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 181
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 YSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKELDR 840

Qy        182 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 241
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 IIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAEIIS 900

Qy        242 VQVPKILSGKVKPIYFHTQ 260
              ||||||||||||||||||
Db        901 VQVPKILSGKVKPIYFHTQ 919
```

Search completed: October 9, 2003, 11:41:13
Job time : 46 secs